UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Escola Tècnica Superior d'Enginyeria
de Telecomunicació de Barcelona

Characterization of Channel Selection in LTE-U based on Q-Learning

A Master's Thesis

Submitted to the Faculty of the

Escola Tècnica d'Enginyeria de Telecomunicació de Barcelona

Universitat Politècnica de Catalunya

by

Massimiliano Concetti

In partial fulfilment

of the requirements for the degree of

MASTER IN TELECOMMUNICATIONS ENGINEERING

Advisor: Jordi Pérez-Romero

Barcelona, February 2016

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

telecom BCN

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

telecom BCN

Title of the thesis: Characterization of Channel Selection in LTE-U based on Q-Learning

Author: Massimiliano Concetti

Advisor: Jordi Pérez-Romero

Abstract

The usage of LTE in unlicensed bands is a topic of great interest among the telecommunications companies and organizations, given the many advantages it can bring. For this reason, many feasibility studies have been conducted in order to define a standard for this technology. One of the approaches proposed is to apply the Q-Learning algorithm to the Channel Selection in LTE-U. Many promising results have already been achieved, so this technique deserves further insights. The objective of this Thesis is to assess the performance of this algorithm under different conditions of parameter settings and scenarios in order to optimize the performance of the technique by modifying the parameters and the operation procedure. The items analysed are the impact of the Initial Temperature, of the Learning Rate, of the positions of the SCs and of the number of active users.

Table of contents

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

telecom BCN

List of Figures

List of Tables

1. Introduction

Mobile data exchange over the cellular network has seen an impressive growth in the last years as a result of the enormous number of smartphones and tablets sold. Forecasts about the future of telecommunications say that mobile data traffic will be increasing more in the next years as a consequence of smart devices, smart cities, machine-to-machine (M2M) communications and more.

LTE Unlicensed (LTE-U) is an improvement of LTE Advanced which extends LTE Carrier Aggregation (CA) from licensed spectrum to unlicensed spectrum using LTE small cells to allow opportunistic data offload and to offer the best mobile broadband to users.

The use of unlicensed spectrum is an important enhancement to satisfy the always growing demand of ultra-high capacity, although licensed spectrum remains the top priority for operators to deliver advanced services and to match the Quality of Service requested.

Licensed spectrum allows higher performances and better user experience, but operators have to pay big amounts of money to have its exclusive usage. On the contrary, the access to unlicensed spectrum does not require a payment, but its users must not interfere with other users. LTE-U must support fair access to multiple LTE-U and Wi-Fi networks, adapting itself to the presence of other users without degrading their performances.

An approach to the Channel Selection functionality in the unlicensed band for LTE-U cells is through a distributed Q-Learning mechanism that exploits prior experience to decide the most appropriate channel.

Q-Learning is one of the most known Reinforcement Learning (RL) algorithms whose objective is to make a learning system adapt to the surrounding environment by finding an optimal action-selection policy. The system learns an action-value function that gives the expected utility of taking a given action in a given state. The iterative process updates and corrects the expected value every time an action is made, so that the agent, in this case the small cell, learns the optimal configuration to get the best total reward.

A team of professors from the Department of Signal Theory and Communications of the *Universitat Politècnica de Catalunya* (UPC) has published some technical papers about this subject, such as [1] [2] [3].

The promising results presented in those works are obtained from a piece of software they have developed in MATLAB in the past years. The aim of this Master Thesis is to assess the behaviour of the Q-Learning algorithm under different conditions when it is applied to LTE-U, by modifying some key parameters and some elements of the scenario in order to evaluate their influence on the overall performance.

This document has the following structure. Chapter 2 gives an overview of the LTE technology, describing the enhancements from the previous technology and the evolution of the architecture. Besides, the standard LTE Advanced, that is an enhancement of LTE, is also described, with a particular focus on the Carrier Aggregation feature which is relevant for LTE-U.

Chapter 3 introduces LTE-U and discusses about the use of unlicensed spectrum. The Licensed Assisted Access feature is explained, as well as the Channel Selection mechanism in LTE-U.

Chapter 4 provides an overview of the Reinforcement Learning method, introducing the basic principles and presenting the Q-Learning algorithm, in particular when it is applied to LTE-U.

Chapter 5 discusses the evaluation methodology of this Thesis, presenting the model adopted for the throughput characterization and analysing the software used for the simulations.

Chapter 6 presents the studies conducted for this project and the results obtained, i.e. the impact of the Initial Temperature, of the Learning Rate, of the considered scenario, and of the number of active users on the performances of the Q-Learning algorithm.

Chapter 7 provides a cost estimation of this project.

Chapter 8 contains the conclusions and the future development of this work.

The research project for this Master Thesis has been done at the *Universitat Politècnica de Catalunya* (UPC) from September 2015 to January 2016. Table 1 contains the Gantt diagram that shows the schedule for the different tasks of this work.

Activity / Time	September 2015				October 2015				November 2015				December 2015				January 2016			
State of the Art																				
Familiarization with the code																				
Study on the Initial Temperature																				
Study on the Learning Rate																				
Study on the modifications of the scenario																				
Writing of the final report																				

Table 1 Gantt diagram of the project schedule

2. LTE

LTE (Long Term Evolution) is a standard for wireless communications of high-speed data introduced in the Release 8 of 3GPP and it represents the transition from 3G to 4G.

The Third Generation Partnership Project (3GPP) is a collaboration between seven telecommunications standard development organizations from Asia, Europe and North America. The original scope of 3GPP in 1998 was to produce specifications for a 3G mobile phone system based on the evolved GSM core networks. Subsequently, the scope was extended to include the development and maintenance of the standards related to different generations of mobile telecommunications technologies, including LTE [4].

After the immense success of GSM/UMTS standards, 3GPP has delineated the long-term evolution of 3G to ensure the continuity of competitiveness of those technologies for the future [5]. The intention of LTE is to promote the wide band usage in mobility, exploiting the experience and the investments made for 3G networks in order to achieve enhanced performances. LTE offers higher data rate with reduced latency on both user plane and control plane, efficient spectrum utilization and flexible spectrum allocation, improved system capacity and coverage, reduced cost for the operator.

Fig. 2.1 LTE Official Logo [5]

2.1. Motivations for LTE

The huge growth of mobile data in the last years is one of the main reasons for the need of an evolution of the 3G system. For many years, the voice traffic was the leading service in mobile telecommunications networks, but the increased availability of 3.5G communication technologies and the introduction of Apple iPhone and Google's Android operating system, respectively in 2007 and 2008, determined the increase of data traffic. The number of mobile applications has grown exponentially as well as the consumption of mobile data, leading the 2G and 3G networks to a congestion. The success of IP-based services delivered over the Internet, along with the diffusion of location-based services and tracking services, made clear that there would have been a convergence toward the use of internet protocols and that all the future services would have been carried on top of IP. The use of VoIP to transport voice calls over packet switched networks is a clear example of this type of services. Future mobile-communication networks need to be optimized for IP-based applications because operators want to move their business to the packet switched domain. The high data rates requested for the new demanding data applications require delays in transferring data packets between network elements and across the air interface much smaller than the ones introduced in the 3G networks, especially for real-time applications.

The need to maintain the backwards compatibility with earlier devices while adding new features to the system has made the specifications for UMTS always more complex. For all these reasons, when LTE was first standardized, it has been decided to design its radio interface from scratch, improving in this way the performance of the system by optimizing it for IP and avoiding the support to the ISDN traffic.

2.2. Requirements of LTE

The main objectives of the evolution were to further improve service provisioning and reduce user and operator costs. The key performance and the principal requirements were identified at the beginning of the standardization work on LTE in 2004.

The end-user data rate is significantly higher than the previous standards, with target peak of more than 100 Mbps over the downlink and 50 Mbps over the uplink, for 20 MHz spectrum allocation, assuming 2 receive antennas and 1 transmit antenna at the terminal.

The average user throughput per MHz and the spectrum efficiency are 3-4 times better in downlink and 2-3 times in uplink, compared to the 3GPP release 6 [6].

Great improvements are requested for the latency in both user and control plane. In order to improve the performance of higher layer protocols, the time to transmit an IP packet from the terminal to the RAN edge node and vice versa is reduced to less than 30 ms, e.g. sub-5 ms for small IP packets in optimal conditions. On the control plane, the latency for handover and connection setup is lower than the previous radio access technologies, e.g. less than 100 ms to allow fast transition times.

The wide-area coverage is augmented. The possibility for significantly higher data rates over the entire cells includes also a particular attention to the users at the cell edge. LTE supports cell sizes from tens of meters radius up to 100 km radius macro-cells.

LTE bandwidth is not fixed at 5 MHz like it happened in WCDMA, but it supports a subset of bandwidths of 1.4, 3, 5, 10, 15 and 20 MHz. Moreover, LTE is deployed on a per-need basis when and where spectrum can be made available so the LTE radio access is designed to be able to operate in a wide range of frequencies. Technical requirements on the spectrum flexibility have a great importance for the operators because in this way they have the possibility to deploy this system in the existing spectrum they have already paid for. LTE-based radio access can be deployed in both paired and unpaired spectrum because it supports both frequency- and time-division-based duplex arrangements [7]. When using *Frequency Division Duplex* (FDD), the base station and the mobile have simultaneous downlink and uplink on two different carrier frequencies, sufficiently separated from each other. On the contrary, when using *Time Division Duplex* (TDD), the transmission takes place on the same carrier frequency, but in different, non-overlapping time slots. LTE supports them both within a single radio access technology.

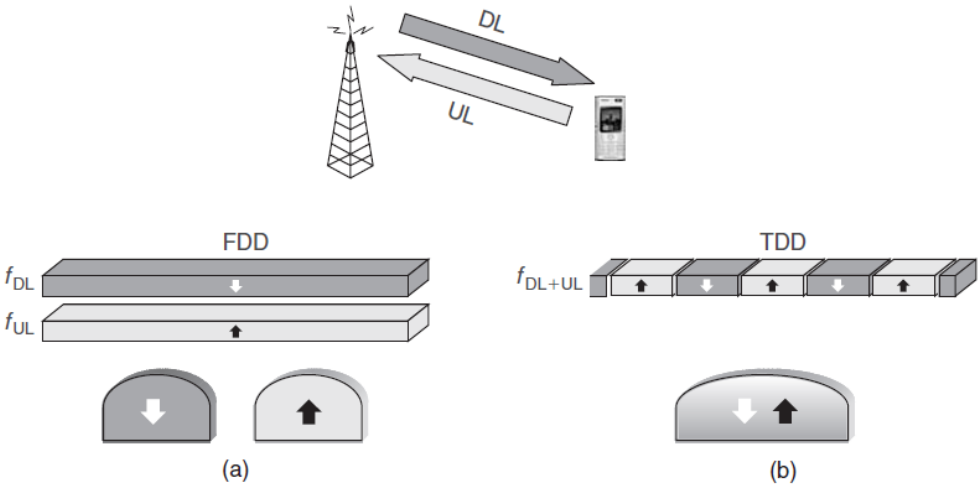


Fig. 2.2 FDD vs. TDD. FDD: Frequency Division Duplex; TDD: Time Division Duplex; DL: Downlink; UL: Uplink [8]

Even if the new LTE architecture is not compatible with the existing UTRAN/GERAN systems and other non-3GPP specified systems, inter-operability with them is ensured. Multimode terminals support handover to and from UTRAN and GERAN, with an interruption time smaller than 300 ms for real time services and smaller than 500 ms for non-real time services.

Not only inter-operability, but also co-existence in the same geographical area and co-location with GERAN/UTRAN is guaranteed. Furthermore, different operators have the ability to use adjacent bands as well as cross-border without interfering with each other.

Operators can maintain their base of end users and deploy the new technology gradually in areas where it is profitable. Service continuity and mobility between systems are two other critical aspects that have been taken into account. The migration to the new technologies was smooth so that operators could reuse sites, investments, and transmission equipment.

LTE supports a communication channel called *Multicast-broadcast single-frequency network* (MBSFN) whose transmission mode is intended to improve the efficiency of the already existing *enhanced Multimedia Broadcast Multicast Service* (eMBMS). This technology can deliver services using the LTE infrastructure without the need for additional expensive licensed spectrum or dedicated infrastructure and end-user devices. The mobile TV is the main service which makes the most of MBSFN and it is a competitor for dedicated mobile TV broadcast systems such as DVB-H. Many TV programs can be broadcasted in a specific radio frequency spectrum as compared to the traditional terrestrial TV broadcasting.

The complexity of terminals and systems was designed to be acceptable, as well as cost and power consumption. All the interfaces specified are open for multi-vendor equipment interoperability and the manufacturers have the ability to reuse investments in development, in order to design and release stable equipment at a competitively price, with a shorter time to market.

The system is optimized for mobility, supporting a wide range of mobile speeds, from the almost stationary scenario where terminals move at 0-15 km/h, up to dynamic scenario of high speed trains, where they move at 350 km/h or 500 km/h depending on the frequency bands.

The end-to-end Quality of Service (QoS) is supported for all the services, also in harsh conditions. For example, the Voice over Internet Protocol (VoIP) is supported with at least as good radio and backhaul efficiency and latency as voice traffic over the UMTS circuit switched networks.

2.3. LTE Architecture

The LTE infrastructure is entirely new and separated from the infrastructure of the previous standards. The requirements for LTE in terms of high throughput, low latency and optimization for packet data pushed the standardization organs to design a new simplified architecture with fewer restrictions on backwards compatibility. The architecture, named *Evolved Packet System* (EPS), is purely IP based and it is composed by three main components: the *user equipment* (UE), the *Evolved UMTS Terrestrial Radio Access Network* (E-UTRAN) and the *Evolved Packet Core* (EPC). The radio interface E-UTRAN is the combination of the new LTE air interface (E-UTRA) and a network of base stations, called *eNodeBs* (eNBs). Fig. 2.3 shows the evolution of the network, from GSM to LTE.

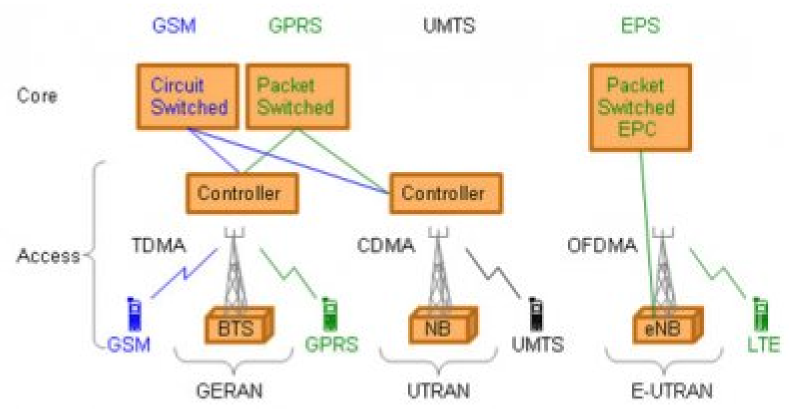


Fig. 2.3 Network solutions from GSM to LTE [5]

Five different LTE user equipment categories were defined in the 3GPP release 8, depending on the maximum peak data rate and MIMO capabilities support. The eNodeBs are the evolution of the nodeBs present in the UTRAN of UMTS. The NBs had just a minimum functionality because the radio resource management was made by the controller, called *Radio Network Controller* (RNC). The architecture of E-UTRAN instead is simplified because there is no separate control element and all the functionalities are executed by the eNodeBs which generate a flat architecture, as opposite to the hierarchical architecture of the previous systems. The eNodeBs are interconnected via the X2 interface and towards the core network by the S1 interface, as it is shown in Fig. 2.4. The connection set-up time and the time required for a handover are reduced in this distributed system because the mobile has to exchange information just with its eNB, which is coordinated with the neighbouring eNBs.

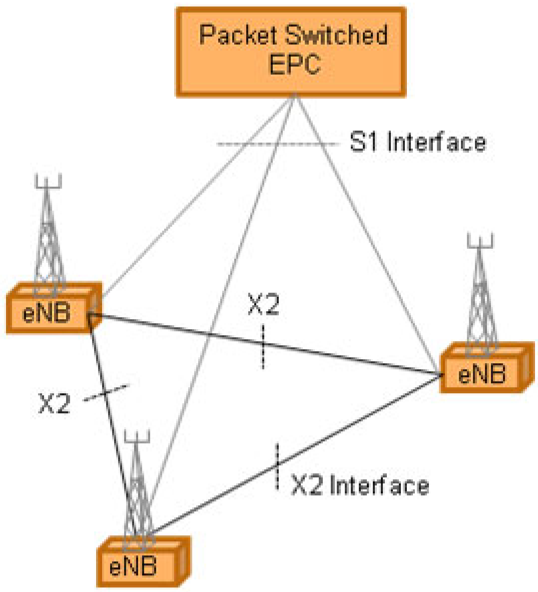


Fig. 2.4 X2 and S1 Interfaces [5]

The air interface E-UTRA is the radio-based communication link between the user equipment and the eNodeB. It uses *Orthogonal Frequency-Division Multiplexing* (OFDM) and *Multiple-Input Multiple-Output* (MIMO) antenna technology, depending on the terminal category.

The first one is a method of encoding digital data on multiple carrier frequencies, known since the mid-1960s but too complicated or expensive to be implemented with the technology of that time. It has now become a popular scheme for wideband digital communication, used in a number of non-cellular wireless systems such as *Digital Video Broadcasting* (DVB). The basic idea behind OFDM is to use many orthogonal narrowband sub-carrier signals to carry data on several parallel channels instead of using a single wideband carrier. The subcarriers are closely spaced and each of them is independently modulated at a low data rate with a conventional modulation scheme, like QPSK, 16QAM, and 64QAM. The total data rate obtained from all the subcarriers is similar to the one obtained by a single-carrier modulation scheme in the same bandwidth. Though, OFDM is much more resistant to the damaging effects of multipath delay spread (fading) in the radio channel. The low symbol rate permits the use of a guard interval, known as cyclic prefix, between each transmitted data symbols, making it possible to eliminate the *inter-symbol interference* (ISI). The OFDM transmitter is typically implemented using low-complexity *inverse fast Fourier transform* (IFFT). In this way the receiver can perfectly detect the transmitted signal provided using the *fast Fourier transform* (FFT), if the maximum delay spread in the channel is shorter than the length of the cyclic prefix.

In downlink LTE uses the *Orthogonal Frequency-Division Multiplexing Access* (OFDMA) which is a multi-user version of OFDM that increases system flexibility by multiplexing multiple users onto the same subcarriers. Therefore simultaneous low data rate transmissions from several users are possible.

In uplink a *Single-Carrier FDMA* (SC-FDMA) scheme is adopted. The transmission processing is similar to the OFDMA, but in this case there is an additional DFT processing step preceding the conventional OFDMA processing. This hybrid modulation scheme combines the low peak-to-average ratio of traditional single-carrier formats with the multipath resistance and frequency scheduling flexibility of OFDM [9]. Fig. 2.5 compares the two technologies, showing as an example how they transmit a sequence of QPSK data symbols.

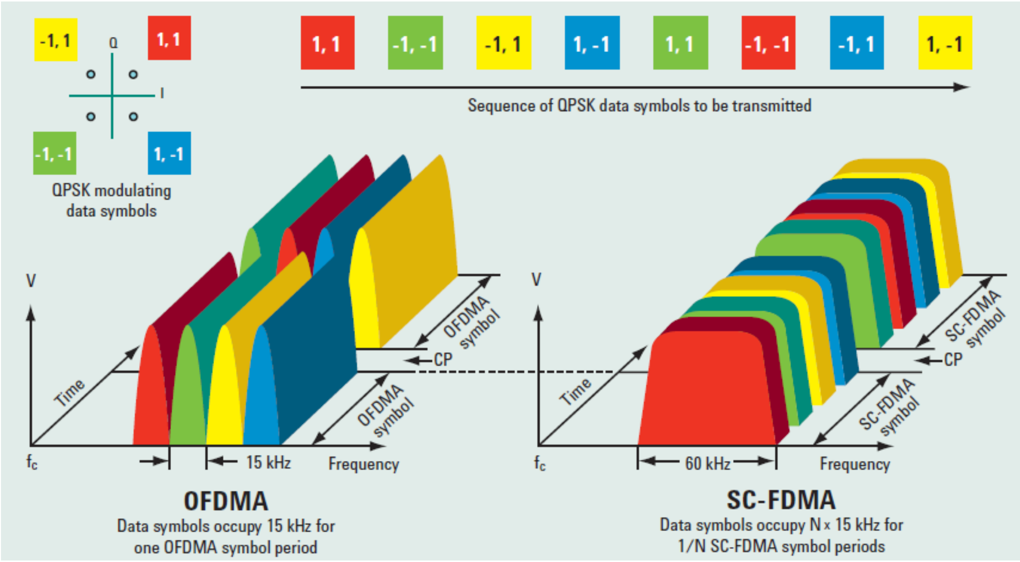


Fig. 2.5 OFDMA vs SC-FDMA [9]

Multiple-Input Multiple-Output (MIMO) is a technique used to increase the peak data rates through multi-stream transmission. It refers to the use of multiple antennas at transmitter and receiver side in order to exploit the multipath propagation. The presence of buildings and other objects in the scenario causes the reflection of the signals, originating multipath fading. MIMO takes advantage of this situation because, using multiple antennas, it can distinguish the different signals received from the different paths and combine them together. The diversity gain permits to achieve high data rates, but a high carrier-to-interference ratio at the receiver is required, so MIMO is mainly applicable in smaller cells or close to the base station, where usually there are higher carrier-to-interference ratios.

The baseline configuration for LTE downlink is a 2x2 MIMO, i.e. two transmit antennas at the base station and two receive antennas at the terminal side. While the number of antennas at the base station can be increased without major difficulties, on the terminal side this number is limited by the dimension of the antennas.

Three main different multi-antenna configurations are possible for MIMO: *diversity*, *spatial multiplexing*, and *beamforming*,

Diversity has been used since the early days of mobile communications to exploit diversity and increase the robustness of data transmission. A single stream is transmitted from each of the transmit antenna, so the receiver gets replicas of the same signal. The channels experienced by the different antennas should have low mutual correlation, which can be obtained with a sufficiently large inter-antenna distance. Usually the signal is coded before the transmission with full or near orthogonal coding in order to increase the diversity effect. The signals reach the receivers with different phase shifts that can be easily removed and they are affected by independent fading, if the receiving antennas are far enough apart. Therefore, after the combination of the different received signals, the amount of fading in the resulting signal is reduced, as well as the error rate. Fig. 2.6 illustrates the benefits of using diversity.

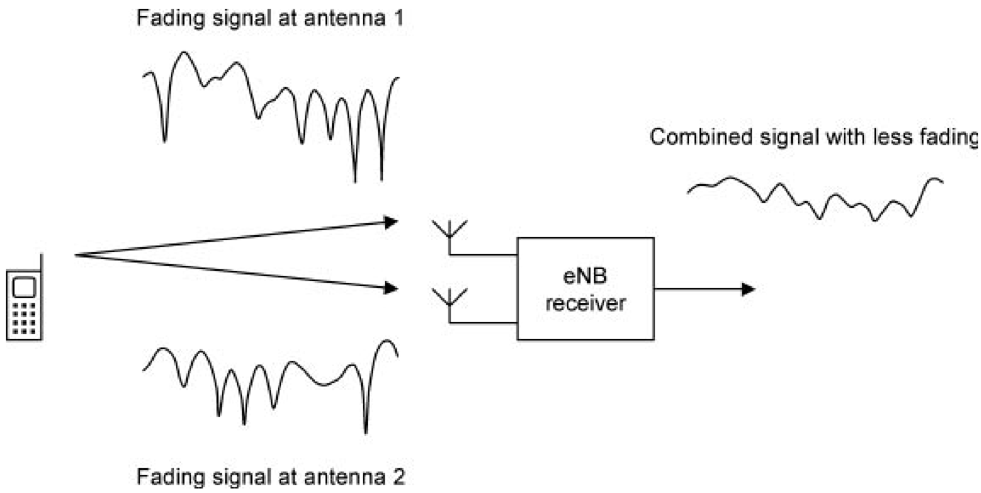


Fig. 2.6 Reduction in fading by the use of a diversity receiver [10]

In spatial multiplexing, the transmitter and receiver both use multiple antennas so as to increase the capacity at higher *signal-to-noise ratios* (SNR). A high-rate signal is split into multiple lower-rate parallel streams and each of them is transmitted simultaneously in the same frequency channel from a different transmitting antenna, exploiting the spatial dimension of the radio channel. The peak data rate is given by the minimum between the number of transmitting antennas and the number of receiving antennas. At the receiver, if the signals are sufficiently separated and the channel state information is precise enough, the streams can be correctly separated. Spatial multiplexing allows very high bandwidth utilization without a reduction in power efficiency and it can be used for single user (SU-MIMO) and multi user (MU-MIMO) transmissions.

Beamforming consists in shaping the overall antenna beam using multiple antennas at the transmitter and/or receiver to maximize the signal power at the receiver. The signals are emitted from different antennas with appropriate phase and gain and they add up constructively, increasing the received signal gain and reducing the multipath fading effect. Fig. 2.7 shows the basic principles of beamforming.

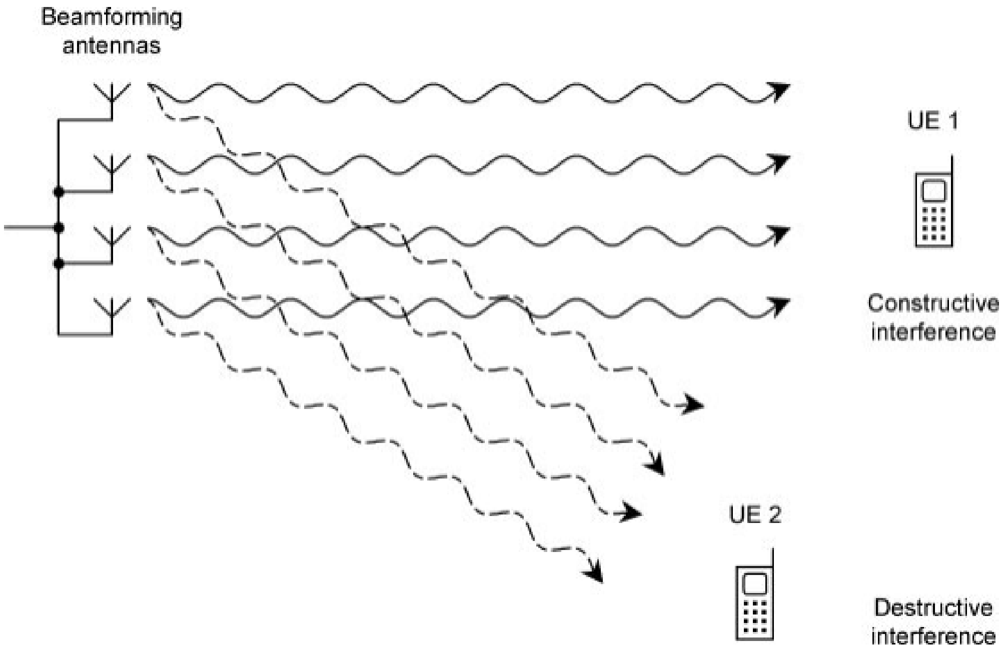


Fig. 2.7 Beamforming [10]

System Architecture Evolution (SAE) is the name used to refer to the 3GPP's work item on the evolution of the GPRS core network. Differently from that, SAE has a flat, all-IP, simplified architecture that supports higher throughput and lower latency radio access networks. The main component of the SAE architecture is the *Evolved Packet Core* (EPC), also known as SAE Core. The requirements for SAE are mainly non-radio access related and they cover high-level user and operational aspects, basic capabilities, multi-access and seamless mobility, man-machine interface aspects, performance requirements for the evolved 3GPP system, and security and privacy aspects. SAE supports the mobility between heterogeneous RANs, including the air interfaces E-UTRA from LTE, GERAN and UTRAN from GPRS and UMTS respectively, and also non-3GPP systems, such as Wi-Fi or WiMAX.

2.4. LTE Advanced

LTE Advanced is a mobile communication standard, defined in 3GPP's release 10 in March 2011. It is a major enhancement of the Long Term Evolution standard and it focuses on higher capacity. LTE Advanced is the real 4G because it provides higher bitrates in a cost efficient way and it completely fulfils the requirements set by *International Telecommunications Union* (ITU) [11]. In 2008 the ITU introduced the term IMT Advanced to identify mobile systems whose capabilities go beyond those of 3G systems, defined in IMT 2000.

The peak data rate can theoretically reach 3Gbps in downlink and 1.5 Gbps in uplink, while in LTE it was just 300 Mbps in downlink and 75 Mbps in uplink [11]. The transmission bandwidth is wider than approximately 70 MHz in DL and 40 MHz in UL. The latency is further reduced, allowing the switch on the C-plane from Idle to Connected in less than 50 ms. The spectral efficiency is increased to a maximum of 30 bps/Hz in downlink and 15 bps/Hz in uplink. The performance at the cell edge is improved, reaching a throughput twice that of LTE, while the average user throughput is instead 3 times. The number of active users per cell that LTE Advance supports in a 5 MHz bandwidth is 300, while in LTE it is 200. The mobility is the same as that in LTE, up to 350 or 500 km/h depending on the frequency band, but system performance is enhanced for slow speeds.

The main new functionalities introduced in LTE Advanced are *Carrier Aggregation* (CA), enhanced use of multi-antenna techniques and support for *Relay Nodes* (RNs).

Carrier Aggregation is a technology which allows increasing capacity by using multiple channels either in the same bands or different areas of the spectrum to provide the required bandwidth. It will be discussed more in detail in the next paragraph, since it is relevant for this Thesis because it is the main concept behind LTE-U and LAA, both descripted in the chapter 3.

The concept of Relay Nodes is introduced in in 3GPP LTE Advance for efficient heterogeneous network planning. The Relay Nodes are low power eNodeBs that provide enhanced coverage and capacity at cell edges [12]. The RNs do not just rebroadcast a signal, but they actually receive, demodulate, and decode the data before re-transmitting a new signal. The eNodeBs connected via the Un radio interface to the Relay Nodes are named *Donor eNodeB* (DeNB). They serve their own UE like a regular eNodeB, but in addition they share their radio resources for RNs. Thanks to the Relay Nodes, LTE Advance can obtain increased coverage and capacity at cell edges as well as at hot-spot areas. Moreover, coverage can be extended in targeted areas at a low cost, for example reaching remote areas without fiber connection. Fig. 2.8 shows a simplified scheme of the connection between a Donor eNodeB and a Relay Node.

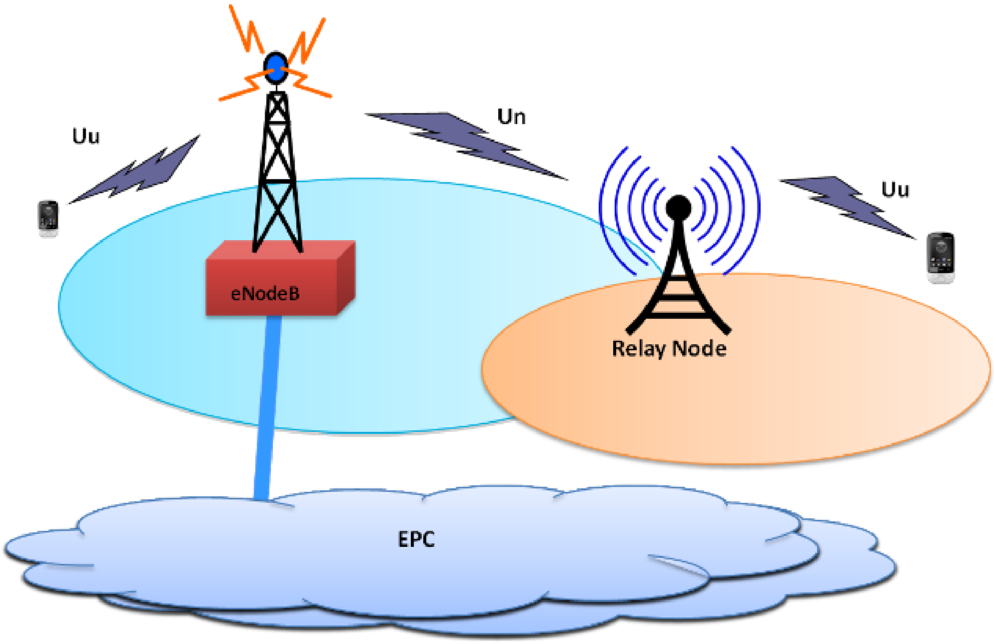


Fig. 2.8 Relay Nodes [12]

2.5. Carrier Aggregation

The Carrier Aggregation (CA) technology allows scalable expansion of effective bandwidth provided to a user terminal through simultaneous utilization of radio resources across multiple carriers. The CA in LTE-Advanced is designed to support aggregation of a variety of different arrangements of component carriers (CCs), including CCs of the same or different bandwidths, contiguous or non- contiguous CCs in the same frequency band, and CCs in different frequency bands [13].

Carrier aggregation is supported by both formats of LTE, namely the FDD and TDD variants. This ensures that both FDD LTE and TDD LTE are able to meet the high data throughput requirements set by ITU.

The component carrier can have a bandwidth of 1.4, 3, 5, 10, 15 or 20 MHz and a maximum of five component carriers can be aggregated, hence the maximum aggregated bandwidth is 100 MHz. In FDD the number of aggregated carriers can be different in DL and UL. However, the number of UL component carriers is always equal to or lower than the number of DL component carriers. The individual component carriers can also be of different bandwidths. For TDD the number of CCs as well as the bandwidths of each CC is normally the same for DL and UL [11].

The principal alternatives for Carrier Aggregation, showed in Fig. 2.9, are the following three:

- ***The Intra-band contiguous carrier aggregation***

This form of CA uses a single band. It is the simplest form of LTE carrier aggregation to implement. Here the carriers are contiguous to each other. The spacing between center frequencies of contiguously aggregated CCs is a multiple of 300 kHz to be compatible with the 100 kHz frequency raster of Release 8/9 and preserving orthogonality of the subcarriers with 15 kHz spacing [13].

- ***The Intra-band Non-contiguous carrier aggregation***

This form is more complicated than the first case where adjacent carriers are used. The multi-carrier signal cannot be treated as a single signal and therefore two transceivers are required. This adds significant complexity, particularly to the UE where space, power and cost are major considerations [13].

- ***The Inter-band contiguous carrier aggregation***

This type of CA uses different bands. It is of exacting use because of the fragmentation of bands - some of which are only 10 MHz wide. For the UE it needs the use of multiple transceivers within the single item, with the usual impact on cost, performance and power. Additional, there are also further complexities resulting from the requirements to reduce inter-modulation and cross modulation from the two transceivers [13].

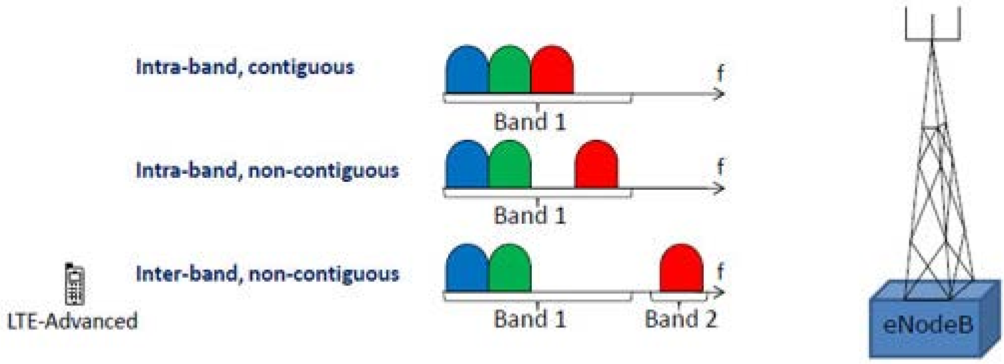


Fig. 2.9 Carrier Aggregation: Inter- and Intra-band alternatives [11]

When carriers are aggregated, each carrier is referred to as a component carrier. There are two categories:

- ***Primary component carrier:*** This is the main carrier in any group. There is a primary downlink carrier and an associated uplink primary component carrier.
- ***Secondary component carrier:*** There may be one or more secondary component carriers.

There is no definition of which carrier should be used as a primary component carrier - different terminals may use different carriers. The configuration of the primary component carrier is terminal specific and will be determined according to the loading on the various carriers as well as other relevant parameters.

In addition to this, the association between the downlink primary carrier and the corresponding uplink primary component carrier is cell specific. Again there are no definitions of how this must be organised. The information is signalled to the terminal of user equipment as part of the overall signalling between the terminal and the base station [14].

3. LTE-U

LTE in Unlicensed Spectrum (LTE-U) is an innovative approach, first proposed by Qualcomm and based on 3GPP Rel. 10/11/12 [15], to allow users to access both licensed and unlicensed spectrum using a single Evolved Packet Core. The main concept of LTE-U is to provide better coverage and larger capacity by aggregating the unlicensed spectrum with the licensed spectrum leveraging the existing Carrier Aggregation technology. LTE-U is considered one of the latest ground-breaking innovations to provide high performance, enhanced user experience, and seamless service continuity between licensed and unlicensed bands under a unified LTE network infrastructure. Operators can offload data traffic onto unlicensed frequencies more efficiently and effectively in order to offer consumers higher data rates, ubiquitous mobility, and improved reliability.

From the operator viewpoint, LTE-U means synchronized and integrated network management, the same authentication procedures, a more efficient resource utilization, and thus lower operational costs.

Radio spectrum is a limited resource and the phenomenal growth of mobile data demands around the world made it scarcer and more valuable. Since it is forecasted that mobile data traffic will increase 1000-fold from 2010 to 2020 [16], it is necessary to optimize the spectrum utilization in a smart and efficient way.

3.1. Initial requirements of LTE-U

The work on LTE-U in the context of 3GPP started within a workshop held in June 2014 in Sophia Antipolis, France. The main ideas that came up are listed in the Workshop & 3GPP TSG-RAN Chairman Dino Flores's summing up [17].

The initial focus should be on *Licensed Assisted Access* (LAA), an operation to aggregate a primary cell, using licensed spectrum, to deliver critical information and guaranteed Quality of Service, and a co-located secondary cell, using unlicensed spectrum, to opportunistically boost data rate.

Although the core technology should be as frequency agnostic as possible in order to be later ported to other frequencies if needed, a clear focus is placed on unlicensed operation in the 5 GHz band [17]. This band is attractive because, even if there may be other operators or systems (e.g. Wi-Fi), there are approximately 500 MHz of spectrum available for various purposes. Regulatory requirements are needed to ensure a fair coexistence between LTE and other technologies or other LTE operators.

Despite different regional requirements emerged from the discussion, such as power level and channel sensing, most of the companies prefer 3GPP to work on the standardization of a global solution that can work across regions. The fact that the 5 GHz band contains available spectrum in many countries of the world is another motivation to put the initial focus on these frequencies.

There is a strong interest to study both indoor and outdoor deployments, using different models and the corresponding modes of operation.

3.2. Advantages and disadvantages of Unlicensed Spectrum

Licensed spectrum provides secure, reliable, and predictable performance so it has been operators' first choice and it contributed to determine the incredible success of mobile networks. Nevertheless, it is costly and limited in availability so it is easily congested and it requires sophisticated inter-cell interference management, especially after the deployment of a large number of small cells.

Hence, operators are motivated to exploit the readily available unlicensed spectrum to fulfil the always growing need of more bandwidth.

The unlicensed spectrum has enabled many low-cost wireless services, including Wi-Fi. Telecommunications organs such as the *Federal Communications Commission* (FCC) in the USA have released several bands for unlicensed commercial use, like the *Industrial, Scientific, and Medical* (ISM) at 2.4 GHz. This band is currently the most utilized band shared by different wireless users such as cordless phones, RFID, ZigBee, Bluetooth, and Wi-Fi enabled device. On the contrary the 5 GHz band is less congested and mainly used by Wi-Fi devices so it is more appealing to develop new services. Moreover, it has a wider available bandwidth, even if it has a shorter communication range which makes it more feasible for transmission of the small cells instead of macro-cells.

The counterpart to the use of available, unlicensed, free-to-use spectrum is finding a compromise between the Quality of Service requested and the fairness towards other systems. There are some fundamental principles and regulations imposed to the technologies working in unlicensed bands to guarantee harmonious coexistence. The first of all is about the transmission power which has to be limited in order to manage the interference among unlicensed users. The *Transmit Power Control* (TPC) is a powerful mechanism used to reduce the power of a radio transmitter to the minimum necessary and, as a consequence, avoid interference and extend battery life [18].

Another technique widely used in unlicensed spectrum is called *Listen-Before-Talk* (LBT) or *Clear Channel Assessment* (CCA). A radio transmitter has to sense the channel before transmitting, in order to detect whether the target channel is occupied by other systems at a millisecond scale. A device can transmit only when no ongoing transmission is observed for a specified period. This is really important for the coexistence between LTE and Wi-Fi systems because, while the former is high-interference-resistant thanks to the centralized MAC and powerful coding schemes, the latter adopts a contention-based MAC protocol with a random back-off mechanism that prevents it from transmitting when it detects LTE transmissions.

Besides, negotiation and coordination policies between LTE operators are needed in order to realize efficient inter-operator spectrum sharing.

Compared to the usage of Wi-Fi in unlicensed spectrum, LTE-U offers several features that are attractive to operators: (i) The spectrum efficiency and coverage with LTE is better than with Wi-Fi due to more advanced radio features such as robust FEC (Forward Error Correction), hybrid ARQ (Automatic Repeat request), interference coordination avoidance, etc., (ii) The same RAN (Radio Access Network) can provide LTE data access in licensed and unlicensed spectrum, (iii) A simplified network management and tracking of KPIs (Key Performance Indicators) through a single RAN can be achieved, (iv) Improved network management and load balancing through tighter integration, (v) Instead of continue pursuing LTE - Wi-Fi interworking, LTE-U is well integrated to the existing operator network, thus solving all authentication, Operations and Management (O&M) and QoS issues, (vi) LTE ecosystem kinds of applications (e.g., machine-to-machine, device-to-device, etc.) are exploitable in LTE-U [1].

3.3. Licensed Assisted Access (LAA)

LAA is the key feature that will enable LTE Advance to aggregate unlicensed frequencies to the licensed ones. 3GPP refers to LTE-U as LAA-LTE to reflect the role of licensed spectrum in its operation because, at least in this initial phase, there is no intention to support standalone operation of LTE in unlicensed spectrum.

The fundamental principle of LAA is that the services offered by LTE Advance are provided through the licensed spectrum, while the unlicensed frequencies are opportunistically used to boost throughput and capacity [19]. There are two modes LAA can be used, the principal one is as a *supplementary downlink* (SDL) data channel. Like the name suggests, it is used only for downlink since the amount of data traffic in this direction is significantly larger than the uplink data traffic. The second mode is as a TDD data channel, for both downlink and uplink, even if it presents more onerous requirements on mobile devices.

The principle of LAA is showed in Fig. 3.1. A mobile terminal is connected to a *primary cell* (PCell) in the licensed bands and one or more *secondary cell* (SCell) in the unlicensed spectrum. The primary cell has the same tasks it would have with LTE Advance, such as providing a robust connection for control signalling, mobility and user data. The secondary cell instead uses unlicensed spectrum to boost the throughput carrying user data in a best effort way, depending on the availability of free channels. Users are anchored to the use of licensed frequencies because they can get access to the SCells only through the PCell.

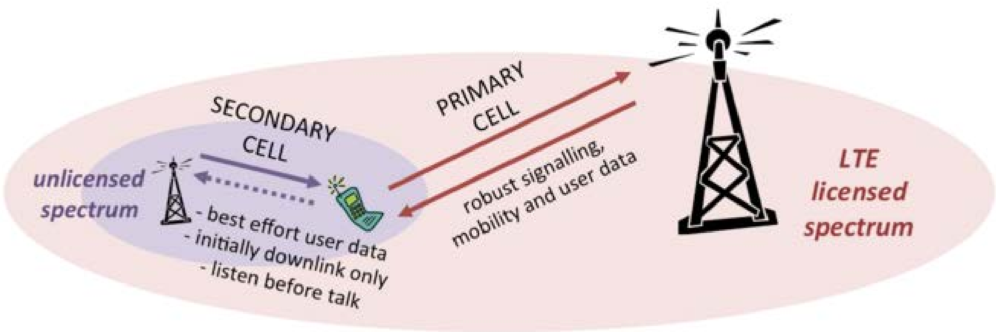


Fig. 3.1 Principles of LTE Licensed Assisted Access [19]

While the expectations for LTE-U are high, this technology has to face many challenges before being brought to fruition. Its success depends heavily on the fair co-existence with other wireless systems, mainly Wi-Fi which is widely used in the 5 GHz band. Part of the enormous success of Wi-Fi can be largely attributed to its capacity to co-exist and share the unlicensed spectrum fairly. There are a number of methods adopted by Wi-Fi based systems to ensure fairness, such as *Listen Before Talk* (LBT), *Channel Selection* (CS), *Physical Carrier Sensing* (PCS), *Virtual Carrier Sensing* (VCS), *Random Back-Off* (RB), *Discontinuous Transmission* (DTX) [20].

The idea of 3GPP about fairness is that LAA should not impact Wi-Fi services (data video and voice services) more than an additional Wi-Fi network on the same carrier. For this reason, LBT is a mandatory requirement in LAA because it is one of the main mechanisms for achieving a fair coexistence. In principle, a transmitter has to apply a *Clear Channel Assessment* (CCA) to make sure that the channel is available before transmitting. The presence or absence of other signals on a channel is determined by the energy detection or preamble detection and the decision is made according to the results of these techniques.

Many times the terms LTE-U and LAA are inappropriately used like synonyms but, even if the two technologies are very similar, there are some differences between them.

The term LTE-U refers in general to the usage of the unlicensed spectrum to boost the capacity and the data rate of LTE Advanced users, while the term LAA indicates the effort of 3GPP to standardise operation of LTE in the Wi-Fi bands.

The development of the LTE-U standard is made by the *LTE-U Forum*, a consortium formed in 2014 by Verizon in cooperation with Alcatel-Lucent, Ericsson, Qualcomm Technologies, and Samsung. The forum collaborated together and generated the technical specifications that include: minimum performance specifications for operating LTE-U base stations and consumer devices on unlicensed frequencies in the 5 GHz band, and coexistence specifications. The specifications support LTE operation in the 5 GHz UNII-1 and UNII-3 bands as Supplemental Downlink (SDL) carriers, in conjunction with an LTE deployment in licensed bands, based on 3GPP already published Release 10 and later specifications [21].

LAA is defined in 3GPP Rel. 13 as a result of the "Study on Licensed-Assisted Access Using LTE" [22] whose objectives include:

- the definition of an evaluation methodology,
- the possible scenarios for LTE deployments focusing on LTE carrier aggregation,
- the documentation of the relevant requirements and design targets for unlicensed spectrum deployment,
- the identification and evaluation of the physical layer options and enhancements to LTE to meet the requirements and targets for unlicensed spectrum deployments,
- the identification and evaluation of any enhancements needed to LTE RAN protocols and an assessment of the feasibility of base station and terminal operation in the 5 GHz band in conjunction with relevant licensed frequency bands.

The results and findings of this study are documented in the technical report TR 36.889 V13.0.0 [22].

Both the technologies dynamically select the unused channel with the least interference to protect and coexist well with Wi-Fi but LTE-U uses *Carrier-sensing adaptive transmission* (CSAT) to sense other users, while LAA abides by a region-specific LBT policy to sense channel availability and adjust on/off LTE cycling.

For this reason, LAA is meant for mobile operator deployments in Europe, Japan and other regions where LBT is a regulatory requirement for the usage of unlicensed bands. On the contrary, LTE-U is deployed in countries like USA, Korea, India, etc. where other coexistence mechanism can be used [15].

3.4. Channel Selection

Channel Selection (also denoted as carrier selection) is the mechanism used to decide the operating channel (i.e. center frequency and associated bandwidth) where a small cell sets up a LTE-U carrier. Therefore, it can be used as a frequency-domain coexistence mechanism to safeguard that LTE is a "good neighbour" in unlicensed bands without requiring modifications in LTE PHY/ MAC standards, e.g. just by enabling small cells to choose the cleanest channel based on received power measurements. If interference is found in the operating channel and there is another cleaner channel available, the transmission can be switched to the new channel using LTE Rel. 10/11 procedures. This ensures that the interference is avoided between the small cell and its neighbouring Wi-Fi devices and/or other LTE-U small cells, provided that there are clean frequencies available. For most Wi-Fi and LTE-U small cell deployments, Channel Selection is usually sufficient to achieve “good neighbour” coexistence [1].

The channel selection for a given small cell should be able to dynamically identify and capture the relevant context information about the current status of utilization of the candidate channels so that the most adequate ones can be selected. Consequently, smart solutions able to identify the best channels under each specific condition are of high interest for the materialization of all the potentials that LTE-U offers [1].

From an architectural point of view, different approaches for Channel Selection can be envisaged: (a) fully distributed case, where each small cell makes decisions on its own, (b) intra-operator coordination, where decisions for a given small cell take into consideration knowledge about other small cells' configurations belonging to the same operator, (c) inter-operator coordination, where also information about small cells from other operators in the area is available and (d) coordination also with managed Wi-Fis in the area. Notice that unmanaged legacy Wi-Fi unable to explicitly provide information about its configuration may also be present in the scenario. Clearly, higher coordination levels will ease the Channel Selection decision-making. However, higher coordination levels involve more demanding network coordination architectures, information exchange protocols and procedures, etc. [1].

One of the possible techniques to handle the Channel Selection in a fully distributed manner is through the Q-Learning algorithm. A description of its behaviour and of the Reinforcement Learning principles in general, is given in the next chapter.

4. Reinforcement Learning: Q-Learning

Reinforcement Learning is an area of machine learning concerned with how software agents ought to take actions in an environment so as to maximize some notion of cumulative reward. Reinforcement Learning is based on interaction with the surrounding environment, where the learner is not told which actions to take, but instead it has to discover which actions yield the most reward by trying them. Differently from *Supervised Learning*, where a knowledgeable external supervisor teaches the learner providing examples of the desired behaviour, in the Reinforcement Learning an agent must be able to learn from its own experience. It usually goes through the same environment many times in order to learn how to find optimal actions, i.e. actions which offer the highest rewards.

Since there are several actions that may be taken from each state, the agent implements a *state-action value* function, denoted *Q(state, action)*. This value depends on the received reward, on the current reward and on some parameters of the algorithm. There are different updating strategies for this value and one of the most known and widely used is the *Q-Learning*.

In this chapter, after an introduction to the basic principles of the Reinforcement Learning paradigm, the Q-Learning algorithm and its main parameters are presented. The last paragraph explains the approach for the Channel Selection in LTE-U using this algorithm and it illustrates the assumptions that were considered for [1]. The results obtained in that work are the start point of the analyses conducted for this Thesis.

4.1. Basic principles

The agent interacts with the environment at each of a sequence of discrete time steps, $t = 0, 1, 2, 3, \ldots$, and it receives some representation of the environment's state $s_t \in S$, where S is the set of possible states. Depending on the state, the agent selects an action $a_t \in A(s_t)$, where $A(s_t)$ is the set of actions available in state s_t. At the following time step, partially as a consequence of the action taken, the agent receives a numerical reward, $r_{t+1} \in R$, and finds itself in a new state, s_{t+1}. A scheme of the agent-environment interaction is shown in Fig. 4.1 [23].

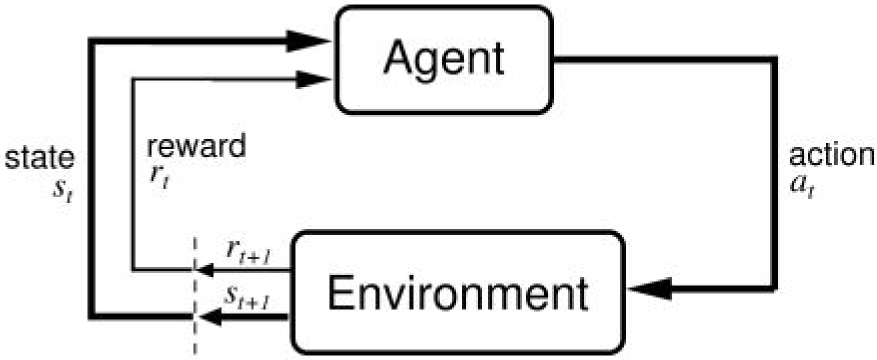


Fig. 4.1 The agent-environment interaction in Reinforcement Learning [23]

The agent learns from past experience and it needs a control policy that will maximize the observed rewards over its lifetime. A policy π_t is the mapping from states to probabilities of selecting each possible action, where $\pi_t(s,a)$ is the probability that $a_t = a$ if $s_t = s$. The policy tells the agent the optimal action to take in any given state with respect to the particular goal the agent wants to achieve.

There are three common policies used for action selection. All of them focus on on-line performance, which involves balancing the exploitation and exploration dilemma in RL-based schemes.

Exploitation means using the information already gathered from the environment to perform the action which brings to the highest reward, according to the past history. On the contrary, exploration is an attempt to discover new features in the environment by selecting a sub-optimal action. Finding the right balance is particularly important to achieve the best performance. Without exploration, the agent will always target the same goal and it will never look for a better one. On the other hand, exploring too much means that the agent discovers the environment, but it does not really learn to take the optimal actions.

The ε-greedy policy makes the agent perform most of the time the greediest action, i.e. the one with the highest estimated reward. The exploration is made every once in a while, selecting with a small probability ε a random action, independently from the value function. If the number of trials are sufficiently long, each action will be tried an infinite number of times, ensuring that optimal actions are discovered.

The ε-soft policy is similar to the previous one because the best action is chosen with probability 1 – ε, while random actions are uniformly chosen in the rest of the time.

The softmax policy is keener than the previous two, because it does not choose an action completely randomly, but it assigns a weight to each of the actions, according to the action-value estimations. In this way the worst actions are unlikely to be chosen, making the exploration phase more effective. The function should tend to choose actions with higher Q-values, but should sometimes select lower Q-value actions. The probability of selecting the highest Q-value action should increase over time. The distribution for this policy, also known as Boltzmann distribution, is as follows:

$$P(a|s) = \frac{e^{\frac{Q(s,a)}{\tau}}}{\sum_j e^{\frac{Q(s,a_j)}{\tau}}} \qquad (4.1)$$

where τ is a positive parameter called *temperature* which controls the probability of selecting non-optimal actions. When the temperature is high, the different actions have almost the same probability to be selected. If τ is close to 0, the best action will be always chosen.

4.2. Q-Learning

Q-Learning is a model-free, off-policy algorithm for *Temporal Difference* (TD) learning. Off-policy algorithms can update the estimated value functions using hypothetical actions, those which have not actually been tried, independently from the policy being followed. In this way the analysis of the algorithm is simplified and early convergence proofs are enabled because an agent may end up learning tactics that it did not necessarily exhibit during the learning phase. The policy is still relevant because it determines which state-action pairs are visited and updated.

The characteristic of being a model-free algorithm is one of the strengths of Q-Learning. The agent is able to compare the expected rewards of the available actions without having any model of the environment, it just needs to know what states exist and what actions are possible in each state.

Each state is assigned an estimated value called Q-value which is updated every time the agent visits that state and receives a reward. At the beginning the value function Q(s,a) is initialized to small random values. The agent observes a state s, picks one of the possible actions according to some policy and then performs it. Depending on the new state s' observed and the reward r obtained, the Q-value is updated according to the general formula defined by:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha \left[r_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t) \right] \quad (4.2)$$

The core of the algorithm is a value iteration update, defined within the *Markov Decision Processes* (MDPs). It assumes the old value and makes a correction based on the new information, weighted by the parameters α_L and γ.

The first one is called *Learning Rate.* It is one of the most important parameters of the Q-Learning algorithm because it determines the impact of the newly acquired information on the future choices the system will do. When it is equal to 0, the agent will not learn anything new, while if it is equal to 1, the agent will take into account only the new information.

The second parameter is the *Discount Factor* and it determines the importance of future rewards. When it assumes the value 0, the agent will consider only the current rewards, the ones which have an immediate effect. On the contrary, when it is equal to 1, the agent will look for long-term higher rewards.

4.3. Q-Learning applied to Channel Selection in LTE-U

The design of a proper Channel Selection functionality can greatly improve the overall efficiency of the LTE-U operation. Therefore, the channel selection for a given small cell should be able to dynamically identify and capture the relevant context information about the current status of utilization of the candidate channels so that the most adequate ones can be selected [1].

There are different possible approaches for Channel Selection. The work made for the development of this Thesis shifts the focus towards a fully decentralized approach in order to understand to what extent the Q-Learning algorithm can overcome the intrinsic disadvantages associated with the fact that no explicit knowledge about the other small cells and/or Wi-Fis operating in the area is available [1].

Each small cell may autonomously learn what channels are usually not being used by its neighbours and then tend to select such free channels. This means that each small cell progressively learns and selects the channels that provide the best performance based on the previous experience [1]. In this context, the *i* small cells are the agents of the Q-Learning algorithm and the *k* channels available correspond to the possible actions that an agent can undertake.

Every time that a decision about a channel is made, the i-th small cells stores a value function Q(i,k) that measures the expected reward that can be achieved by using that channel.

The value function is updated with an easier approach than the general one proposed in [23], because here a null discount rate is assumed. The formula is:

$$Q(i,k) \leftarrow Q(i,k)(1-\alpha_L) + \alpha_L * r(i,k) \qquad (4.3)$$

where $\alpha_L \in (0,1)$ is the learning rate and *r(i,k)* is the reward that has been obtained as a result of the current use of the channel *k*. Assuming that the target of the channel selection is to find a channel that maximizes the total throughput, the reward function considered in this work is given by:

$$r(i,k) = \frac{\overline{R(i,k)}}{R_{max}} \qquad (4.4)$$

where $\overline{R(i,k)}$ is the average throughput that has been obtained by the i-th small cell in channel k as a result of the last selection of this channel. In turn, R_{max} is a normalization factor [1].

At initialization, i.e. when channel *k* has never been used in the past by small cell *i*, *Q(i,k)* is set to an arbitrary value Q_{ini}.

Based on the *Q(i,k)* value functions, the proposed Channel Selection decision-making for the small cell *i* follows the softmax policy in which channel *k* is chosen with probability:

$$\Pr(i,k) = \frac{e^{\frac{Q(i,k)}{\tau(i)}}}{\sum_{k'=1}^{K} e^{\frac{Q(i,k')}{\tau(i)}}} \qquad (4.5)$$

where *τ(i)* is a positive parameter called temperature. High values of temperature cause the different channels to be all nearly equi-probable. Low temperature causes a greater difference in selection probability for channels that differ in their Q(i,k) value estimates, and the higher the value of Q(i,k) the higher the probability of selecting channel k. A cooling function is considered to reduce the value of the temperature *τ(i)* as the number of channel selections made by the small cell increases, so that the amount of exploration will be progressively decreased as the small cell has learnt the best solutions. Specifically, the following logarithmic cooling function is assumed:

$$\tau(i) = \frac{\tau_0}{\log_2(1 + n(i))} \qquad (4.6)$$

where τ_0 is the initial temperature and *n(i)* is the number of channel selections that have been already done by the i-th small cell [1].

5. Evaluation Methodology

This chapter introduces the technical aspects of the research done for this Thesis. Section 5.1 presents the model adopted in this work to assess the throughput that can be obtained in a LTE-U carrier, while the section 5.2 describes the functioning of the software used to assess the performance of the Q-Learning strategy, described in chapter 4. First, the input parameters are presented and then the behaviour of the simulator is explained step by step.

5.1. LTE-U Throughput Characterization

The model for the LTE-U throughput characterization presented here is the same used in [1] and it is based on the same hypothesis.

It is assumed that in the considered scenario, that will be explained in detail in the next chapter, there is a number of small cells denoted as $i = 1,\ldots,S$ making use of the 5 GHz unlicensed band as a supplemental downlink for extending the available capacity in the licensed band. The total band is considered to be organized in channels of bandwidth B, numbered as $k = 1,\ldots,K$.

Considering that the Channel Selection functionality has chosen the k-th channel for carrying out LTE-U transmissions in the downlink of the i-th small cell, and that LBT is required, the total aggregated throughput served by this cell can be estimated as:

$$R(i,k) = \sum_{n=1}^{N(i)} \frac{B}{N(i)} S(SINR_n(i,k)) \frac{1-\theta_{idle}}{M(i,k)} \qquad (5.1)$$

where $N(i)$ is the total number of users being served by the i-th small cell exploiting the supplemental downlink capacity offered by LTE-U; $SINR_n(i,k)$ is the signal to noise and interference ratio observed by the n-th user when downlink data is transmitted on the k-th channel; θ_{idle} is the fraction of time associated with the idle periods imposed by the LBT strategy (CCA time is already included in these idle periods); and $M(i,k)$ is the number of small cells that are sharing in the time domain the k-th channel with the i-th small cell following the LBT strategy (i.e. those that when they are transmitting they are received above threshold TL at the i-th small cell). $S(SINR_n(i,k))$ is a generic function ranging between 0 and S_{max} that provides the spectral efficiency in b/s/Hz as a function of $SINR_n(i,k)$ depending on the characteristics of the technology. In turn, $SINR_n(i,k)$ depends on the propagation conditions between the n-th user and the i-th small cell and on the interference generated by other small cells using the k-th channel and that, when they transmit, they are detected at the i-th cell below threshold TL so that they are not sharing the channel in the time domain based on the LBT [1].

As a result of LBT, expression (5.1) assumes an equal sharing in time domain between small cells, so that on average the i-th small cell can only transmit during a fraction of time $(1 - \theta_{idle}) / M(i,k)$, where θ_{idle} accounts for the waiting periods imposed by LBT and the fraction $1 / M(i,k)$ accounts for the fraction of time that the i-th small cell can transmit because it senses the k-th channel as free. It is also assumed that all the small cells operate only with LTE-U in the downlink direction. Similarly, (5.1) assumes a full buffer traffic model in which the small cell always has data to be transmitted, and that the total bandwidth B is equally shared between all the $N(i)$ users being served by the i-th small cell, so that on average a user observes a fraction $B / N(i)$ of the total bandwidth. It is worth mentioning that expression (5.1) could be easily modified to capture other scheduling strategies to share the bandwidth between users. Note also that (5.1) corresponds to the throughput achievable in one channel. In case that a small cell aggregates multiple channels, the total throughput would be the summation of (5.1) for all the channels [1].

The decision-making applied to perform the channel selection for the i-th small cell will impact on the achieved throughput performances mainly through the terms $M(i,k)$ and $SINR_n(k)$. Thus, if the selected k-th channel is not used by other cells (i.e., $M(i,k) = 1$), higher throughput will follow. Similarly, if the selected k-th channel is affected by low interference levels, high $SINR_n(k)$ will be observed and higher throughputs will follow [1].

5.2. Software

The analysis of the Q-Learning parameters is done by executing many simulations with a software written in Matlab. The software is a simulator for the channel selection in LTE-U and it is suitable for the test with different algorithms besides Q-Learning, such as ITEL (*Iterative Trial and Error Learning*) and its variants, i.e. ITEL-BA and ITEL-BAWII.

Basically, what the program does is to simulate the simple case of two LTE operators trying to gain access to the available channels in an opportunistically way, using one of the algorithms aforementioned. Each of the operators has a definable number of small cells deployed in the given scenario where the users are randomly and uniformly distributed at the beginning of every simulation. The code is explained in the next subparagraphs, the first one illustrating the input parameters and the second one describing the structure of the code.

5.2.1. Input parameters

In the first part of the code, all the input parameters are initialized with the values inserted by the user, according to the type of simulation requested and the specific algorithm that needs to be tested.

The general scenario parameters define the dimension of the scenario and the height of the LTE terminals (UEs) and the height of the small cells (APs). By default it is laid out like the indoor scenario for LTE-U coexistence evaluations defined in [22], as described in paragraph 6.1.

Then, the propagation conditions are specified, firstly defining the frequency and the bandwidth, usually set to 5 GHz and 20 MHz respectively, as suggested in [17] for the initial studies on LTE-U; secondly, characterizing the propagation model and the relating parameters, such as the attenuation and the shadowing, in both *Line of Sight* (LOS) and *Non-Line of Sight* (NLOS) conditions.

After this, the number of access points and users of each operator are defined, as well as their technical parameters, such as the transmit powers and the antenna gains.

Then, there is the definition of the parameters related to the Listen Before Talk (LBT) technique, like the threshold to detect a channel as free, and those related to the throughput computation, like the Signal to Noise and Interference Ratio.

The simulation time is measured relative to a generic unit denoted as “time steps” [1]. The activity time parameter indicates the average number of time steps of the geometrically-distributed activity periods of the small cells, in which they require the activation of a LTE-U carrier to transmit data to their users.

The number of channels available can be specified, as well as their fixed allocation, applicable only in the case of fixed algorithms.

Some binary variables permit to choose if to evaluate the optimum throughput achievable and if to check the convergence behaviour. In case the second one is selected, it is possible to specify the maximum number of realizations of the same experiment; otherwise, this number is automatically set to 1. Besides, the convergence criterion can be set.

Then, it is possible to indicate the algorithms that the two operators are going to use, as well as the specific parameters for each algorithm. In the case of Q-Learning, they are the Initial Temperature τ_0, the Learning Rate α_L, the initial value of the Q-value function, the cooling function, and the reward type.

Lastly, it's possible to define the maximum number of experiments and the duration of each of them, which is determined by the maximum number of time steps.

An experiment simulates the whole process of the Channel Selection, from the deployment of the users in the scenario and the activation of the APs to the evaluation of the performance after the last time step.

A simulation can be composed by multiple independent experiments and its output is given by the average of the results obtained during each of them. A greater number of experiments and time steps improve the quality of the simulation, providing more accurate results at the price of a longer duration so it is important to find an optimal trade-off between them.

5.2.2. Simulation

The main part of the code is entirely enclosed inside a *for cycle* that allows exploring all the values of a given parameter. For each of them, a whole simulation is run and all the results obtained for that specific value are saved in a different file. Right after, there is another *for cycle* that allows making multiple independent experiments for a given simulation. The random number generator at the beginning of this cycle produces a predictable sequence of numbers, starting from a seed inserted by the utilizer. This sequence will be used later for the random deployment of users inside the scenario so, if the seed depends on the specific number of the experiment, every experiment will have the same seed in different simulations.

After this, the Access Points are initialized and they are placed in the scenario, according to the positions described in paragraph 6.1. Furthermore, the program checks the detection conditions between them if they are operating at the same frequency, according to the propagation model and parameters set at the beginning of the code.

Successively, the users are dropped in the scenario, with reference to the random sequence of number generated from the seed inserted. The program consequently computes the distances of each user from the different APs and it associates every UE to the AP that presents the lowest total propagation losses. If a small cell does not have any users associated, it is automatically turned off and put in the inactive state.

Afterwards, the activity is initialized, meaning that the probabilities of initiating and ending a session in the following time step are computed. They depend on the values of the activity time and inactivity time that were defined at the beginning. For example, if the activity time is equal to 1, the probability of ending the session is a certain event so the small cells are going to make a decision about the most suitable channel at every time step.

Then, the software computes all the possible combinations of active small cells and available channels, evaluating also the optimum ones if requested. The optimum combinations are those which maximize the overall throughput and they depend on the detection range of the APs and, consequently, on their reciprocal distances. Obviously, if the number of channels is higher or equal than the number of APs, the optimum is known a priori because it is just a combination where each AP has a different channel.

At this point, the real simulation starts with the execution of a *for cycle* where the software implements the algorithm and makes the computations for each of the time steps designated. Another *for cycle* repeated for every time steps checks all the APs to find the active ones. If an AP is active, which is true when there is at least one user associated to it, the program estimates the *Signal To Noise + Interference Ratio* (SINR) seen by each of the users. The total interference increases if there are other active APs that use the same channel and they are out of the detection range of the current AP. Those small cells create interference because they are received below the threshold TL and so they do not share the channel in the time domain following the LBT strategy. The instantaneous measured throughput for each AP is updated every time that the SINR of a user is higher than a target value.

After this, the software checks again the number of active APs which use the same channel, but this time it just counts those that are detected by the current AP. This number is a parameter in the formula for the computation of the total aggregated throughput showed in the previous paragraph.

Moreover, the average throughput per cell is computed at this point, updating the value after every time step.

Later, the convergence behaviour is checked, if this option was selected before launching the simulation. The software controls if the selection probability of every channel and for all the APs is greater than the convergence criterion that was defined at the beginning. If the system has converged, the result is stored.

Afterwards, the software checks the activity and the channel selection for the next time step. This means that it controls the APs that will be off in the following time step, according to the values selected for the activity and inactivity time. If an AP is ending the session, the reward is updated as well as the value function Q. Otherwise, if the session is not ending in the current time step, the software checks if the quality on the selected channel is sufficiently good or if there is need to change it.

The software also checks all the APs which are going to start a new session in the following time step. For those, it starts the procedure of channel selection based on the algorithm selected.

In the last part of the simulation, the software computes many statistics related to the experiments executed during the simulations, such as the average throughput or the optimum throughput.

Besides, the software gives the possibility to plot some of the results obtained in a graph, such as the evolution of the selection probabilities.

6. Results

Section 6.1 of this chapter describes the default scenario for the some of the analyses performed in this research.

Section 6.2 presents the study about the impact of the Initial Temperature on the average throughput for different combinations of parameters, like the number of channels available or the type of cooling function selected.

Section 6.3 discusses the influence of the Learning Rate, showing the analyses of the convergence behaviour, of the selection probabilities, and of the average throughput.

Section 6.4 presents the effects of some modifications in the scenario, like the positions of the SCs or the number of active UEs.

6.1. Scenario

The considered scenario is based on the indoor scenario for LTE-U coexistence evaluations defined in the context of the corresponding 3GPP Study Item [22]. It consists of a single floor building where two operators deploy 4 small cells (SCs) each. Small cells are equally spaced and centered along the shorter dimension of the building, as depicted in Fig. 6.1.

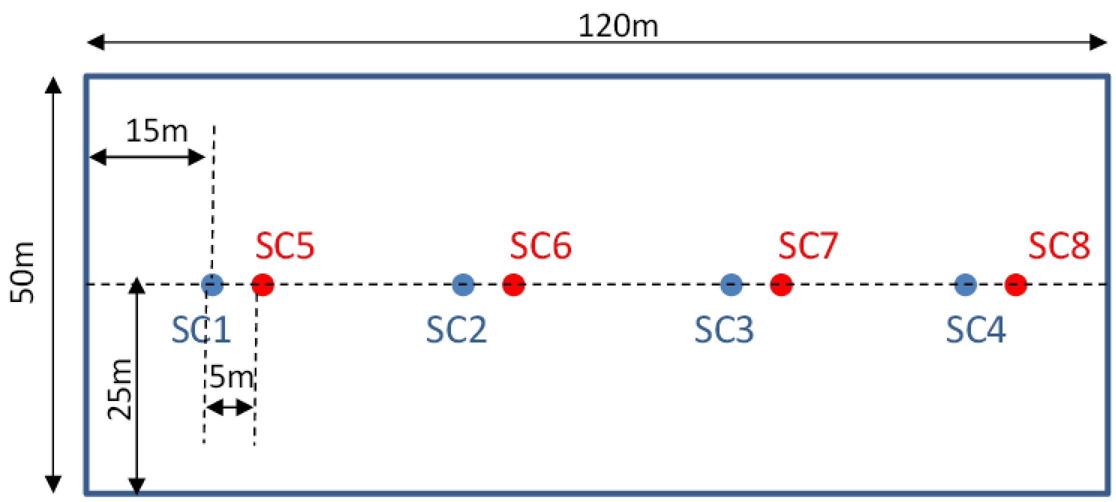


Fig. 6.1 Layout of the floor building

Small cells SC1 to SC4 are owned by operator 1 (OP1) while SC5 to sC8 are owned by operator 2 (OP2). SCs are deployed at height 6m while the antenna height of the mobile terminals is 1.5m. A total of 10 terminals (users) per operator are randomly distributed inside the building. Each user is associated to the SC of its own operator that provides the highest received power. The SC-to-terminal and SC-to-SC path loss and shadowing are computed using the ITU InH model in [24]. The carrier frequency is 5 GHz and the channel bandwidth B ≈ 20 MHz. The transmit power in one LTE-U carrier is 15 dBm. Omnidirectional antenna patterns are assumed with a total antenna gain plus connector loss of 5 dB. The terminal noise figure is 9 dB. The spectrum efficiency function S(SINR) is obtained from Section A.I in [25] with S_{max} = 4.4 b/s/Hz. The threshold TL used in the CCA of the LBT to decide whether a channel is sensed as free or not is TL = -70 dBm/MHz according to the formula in [26]. With this threshold and the considered propagation model, it turns out that in this layout only SC3 and SC6 are able to detect the transmissions of all the other small cells. On the contrary, SC1 is not able to detect the transmissions of SCs 4, 7 and 8; SC2 does not detect SC8; SC4 does not detect SCs 1 and 5; SC5 does not detect SCs 4 and 8; SC7 does not detect SC1; and SC8 does not detect SCs 1, 2 and 5. Besides, θ_{idle} = 0.05 is assumed. The parameters related to the Q-Learning algorithm are α_L = 0.1, τ_0 = 0.15 and Q_{ini} = 0.5 [1].

Given this configuration, the simulator computes the best combination of small cells and channels when all the SCs are active. An optimum solution is a channel assignment that maximizes the aggregate throughput for all the small cells.

6.2. Case study 1: Initial Temperature

The temperature τ is a positive parameter that controls the probability of selecting non-optimal actions and it is included in the formula for the computation of the selection probability. High temperatures make the channels have almost the same probability of being selected, so low temperatures are usually better. The value of the temperature can be fixed and equal to the Initial Temperature τ_0 when no cooling strategy is considered or it can be decreased using a logarithmic cooling function. There are two of them, the first one depends on the number of channel selections already made and is called *cooling* *samples* (6.1), while the second one, called *cooling time* (6.2), is a function of the number of time steps already elapsed.

$$\tau(i) = \frac{\tau_0}{\log_2(1 + n(i))} \qquad (6.1) \qquad\qquad \tau(i) = \frac{\tau_0}{\log_2(1 + t)} \qquad (6.2)$$

Many simulations with different parameter settings have been run to assess the impact of the initial temperature τ_0 on the obtained performance, in particular on the average throughput.

The reference scenario and the parameters of the Q-Learning algorithm are described in the previous paragraph.

Every simulation runs 50 experiments and each of them lasts for 100000 time steps.

6.2.1. Number of channels available k = 4.

The number of channels available has a big impact on the performances of the algorithm. Every small cell tries to have access to a free channel in order to obtain higher throughput and to avoid interference with the other cells. When there are just four channels available, the 8 small cells have to share them and this leads to worse results, compared to the case of 8 channels available.

Basically, the average throughput is inversely proportional to the Initial Temperature. The average throughput obtained at the end of a simulation is a result of how good is the channel available selected by a small cell. The Q-Learning algorithm learns from the past and keeps track of previous choices in a table called Q(i,k) which is constantly updated. The probability of selecting a channel contributes to find the new value for Q(i,k) and it depends on the temperature. The cooling function reduces the value of the initial temperature after every channel selection and provides a new value of temperature. It has to be as low as possible so that the probabilities to choose different channels are more dissimilar. Obviously, the cooling function will need more time to reduce the temperature if the initial temperature is high and this is the reason for the reduced throughput.

The first case presented is a simulation where operator 1 is using the Q-Learning algorithm with the *Cooling Samples* function, while operator 2 has a fixed channel assignment. The results of this simulation are shown in Fig. 6.2 and the behaviour obtained is in line with what was expected: better performances are obtained with lower values of initial temperature.

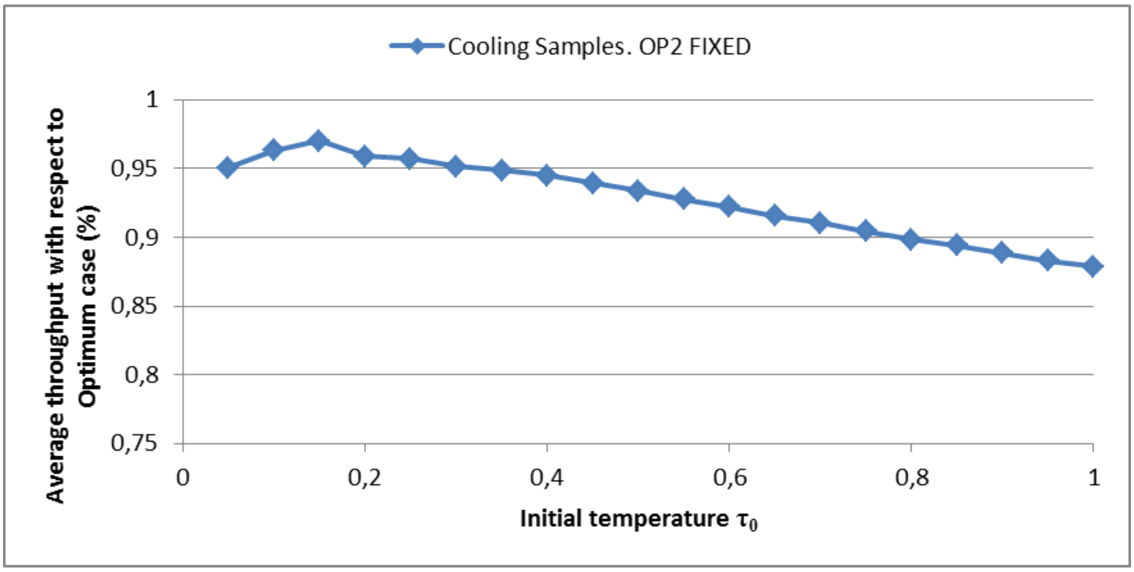


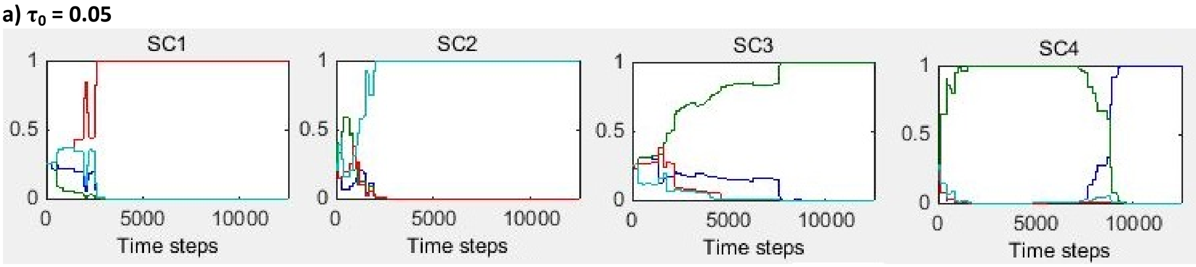
Fig. 6.2 Q-Learning, OP2: FIXED. Cooling Samples

The average throughput starts from 0.95 when the initial temperature is 0.05 and it rises until it reaches the peak when τ_0 is equal to 0.15. After this point, the trend of the curve constantly decreases as long as the initial temperature rises.

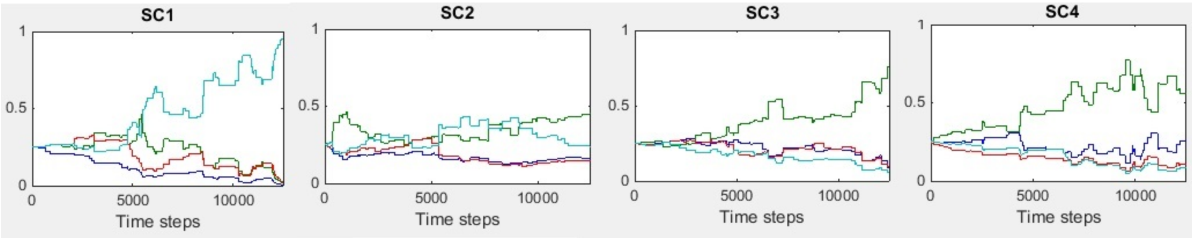
This effect can be seen also in Fig. 6.3 which represents the dynamic evolution of the selection probabilities of every channel from the small cells of operator 1 view point. In the first strip, when the initial temperature is low, the selection probabilities are quite different from each other so the cells can choose the optimal channel and get a better throughput.

The higher values of temperature in the second and the third strip lead to a chaotic situation where the four channels have similar probabilities to be chosen. More time steps are needed to take a decision, causing a decrement of performances.

a) τ_0 = 0.05



b) τ_0 = 0.25



c) τ_0 = 0.65

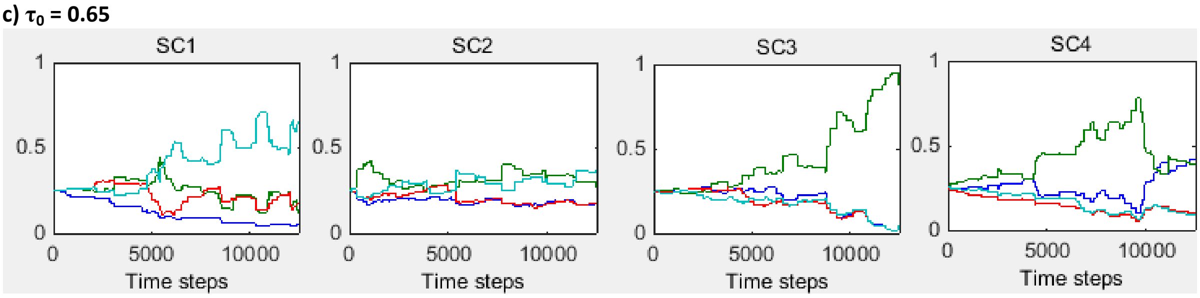


Fig. 6.3 Dynamic evolution of the selection probabilities

In order to understand the importance of the cooling function, it's useful to compare the average throughput obtained when operator 1 uses the cooling samples function and when it uses none.

When the Q-Learning algorithm is applied without a cooling function, temperature is not reduced after the small cell takes a decision, but it is constant to the value of the initial temperature. Performances are worse because the channels are all nearly equi-probable so the optimal one is not always chosen.

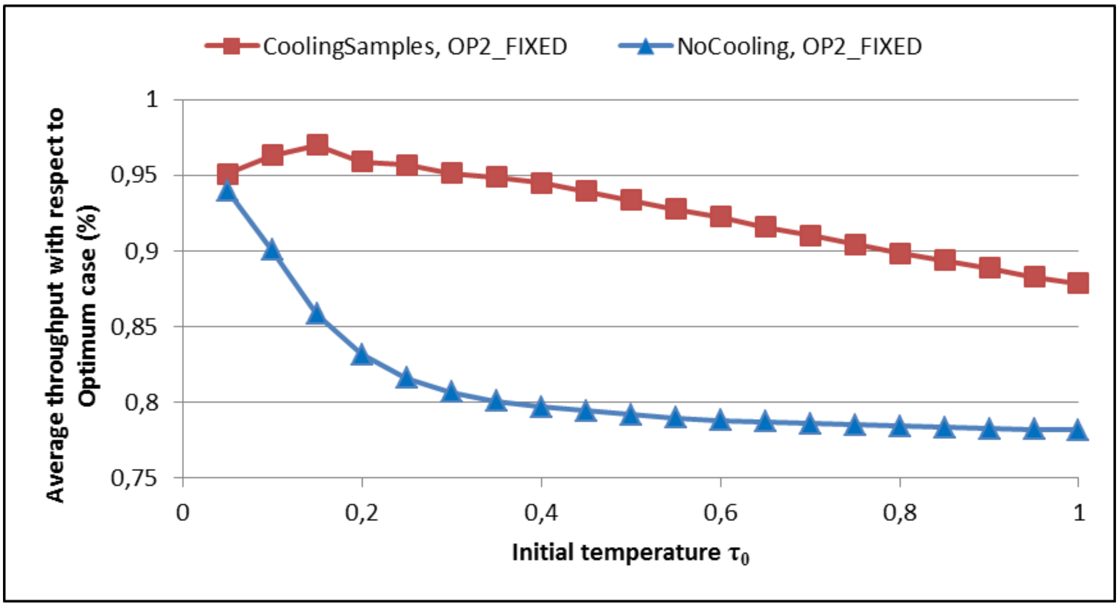


Fig. 6.4 OP1: Q-Learning, OP2: FIXED. No Cooling

The different behaviour of the two curves can be clearly seen in Fig. 6.4. For τ_0 equal to 0.05, the average throughput is the same because the temperature is the lowest possible and it makes no difference reducing it or not.

When the initial temperature increases, the curve corresponding to the no cooling case decreases rapidly until it becomes steady at about 0.78.

On the contrary, when the cooling samples function is applied, the average throughput has a slow constant decline after it reaches a peak in τ_0 equal to 0.15.

An interesting comparison between the Q-Learning with no cooling function and the random channel selection is depicted in Fig. 6.5.

When the cooling function is absent and the temperature assumes large values, the behaviour of this curve approaches that of the random selection case.

The latter case has the lowest performances because the small cells do not chose the channel in the optimal way. Furthermore, since the initial temperature is a parameter of the Q-Learning algorithm, it has no impact on the random selection of the channel, which has a constant average throughput, independent from the temperature.

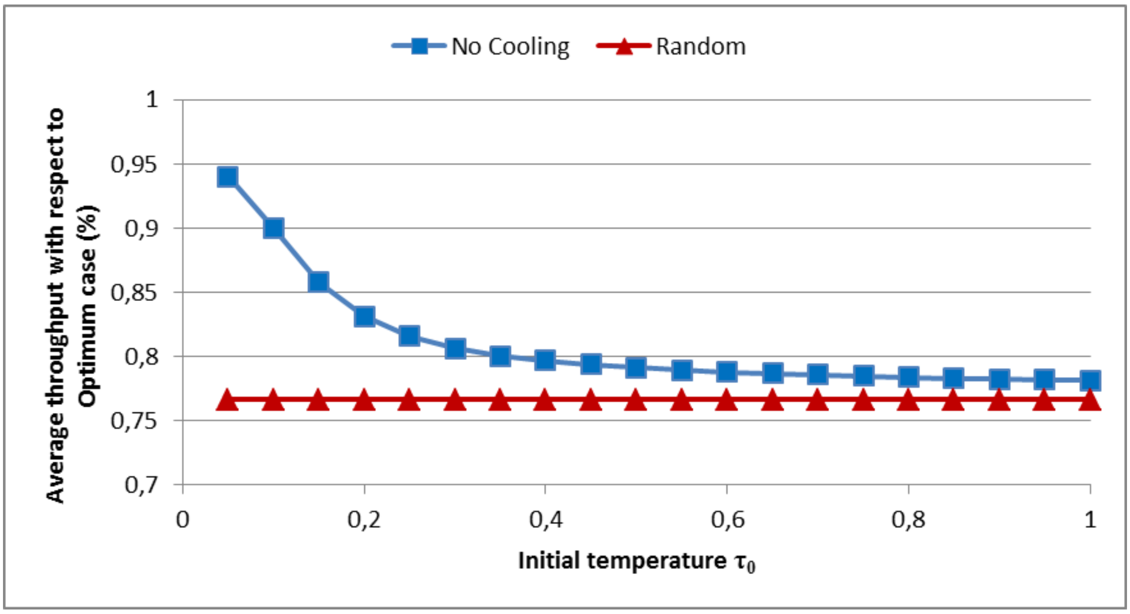


Fig. 6.5 OP1: Q-Learning, OP2: FIXED. Cooling Samples

The second cooling function adopted in the Q-Learning algorithm is called Cooling Time. The effect of using this function is the same as using the Cooling Samples, but, this time, the temperature value is decreased after every time step instead of after every channel selection. In principle, the performances should be better because this function reduces the value of the temperature more frequently than the Cooling Samples, so the difference among selection probabilities should be greater. The curves referred to the two cooling functions are presented in Fig. 6.6.

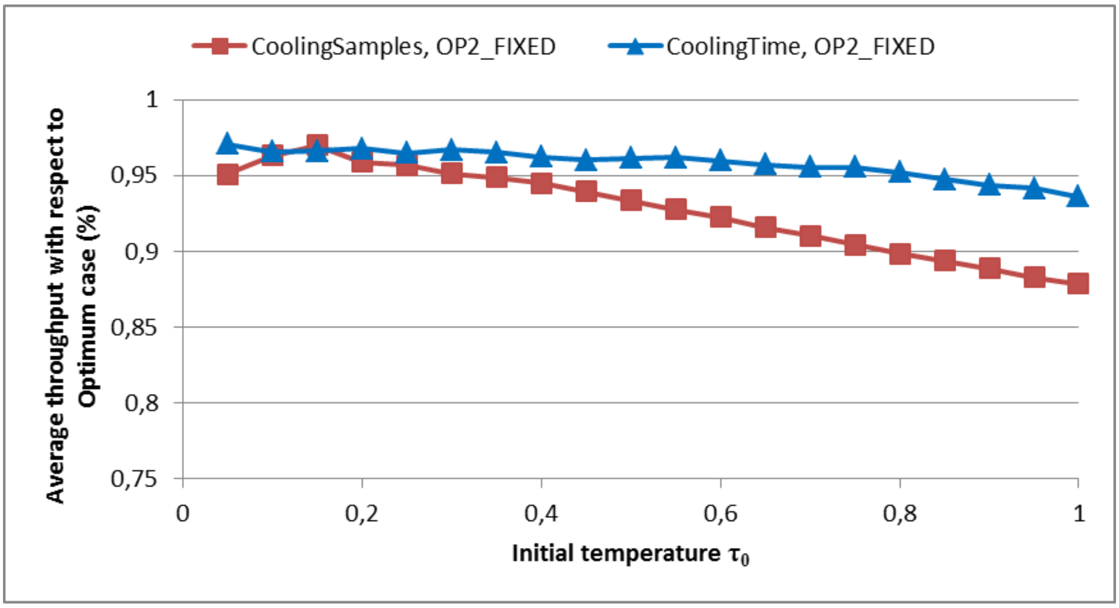


Fig. 6.6 OP1: Q-Learning, OP2: FIXED. Cooling Time

When the Cooling Time is used, the average throughput is almost constant until the initial temperature becomes greater than 0.6 and the curve starts to go down really slowly.

Although the throughput obtained with the Cooling Samples is approximately the same, the curve has a different trend because it starts decreasing sooner and with a bigger slope.

In all the cases presented till now, operator 2 had a fixed assignment. In Fig. 6.7 is given a comparison between the results achieved when only one operator is using Q-Learning and when they both do.

In the former case the 4 small cells of operator 2 have already been assigned the 4 channels available, so the 4 cells left apply the algorithm to find the best way to share them. On the contrary, in the latter case all the 8 small cells are using Q-Learning to decide which channel to use and this bigger exploitation brings to reduced performances.

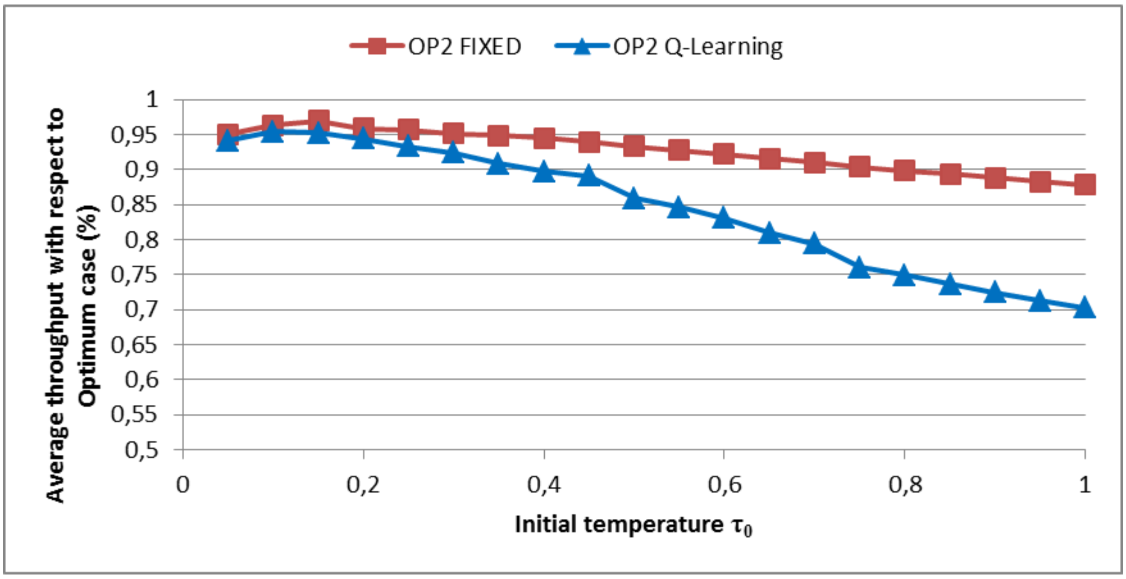


Fig. 6.7 OP1: Q-Learning, OP2: Q-Learning. Cooling Samples

The two curves have the same trend, but the fixed assignment allows better throughput for all the values of the initial temperature, even if there is a slow decline for higher values of temperature.

When both operators apply Q-Learning, the curve goes down much faster as the temperature rises because the decisions about the optimal channel requires more time.

6.2.2. Number of channels available k = 8.

When the number of channels is the same as the number of small cells, the situation is much more favourable. In this case the optimal throughput achievable is 1 because every single small cell can find its own channel and transmit without interfering with the other ones.

At this point, it is convenient repeating the comparison between the cases of operator 2 using a fixed assignment or the Q-Learning, but this time showing the difference when the number of channels changes.

The graph of Fig. 6.8 represents the absolute average throughput instead of the one referenced to the optimal case because, in this case, it offers a clearer observation of the trends.

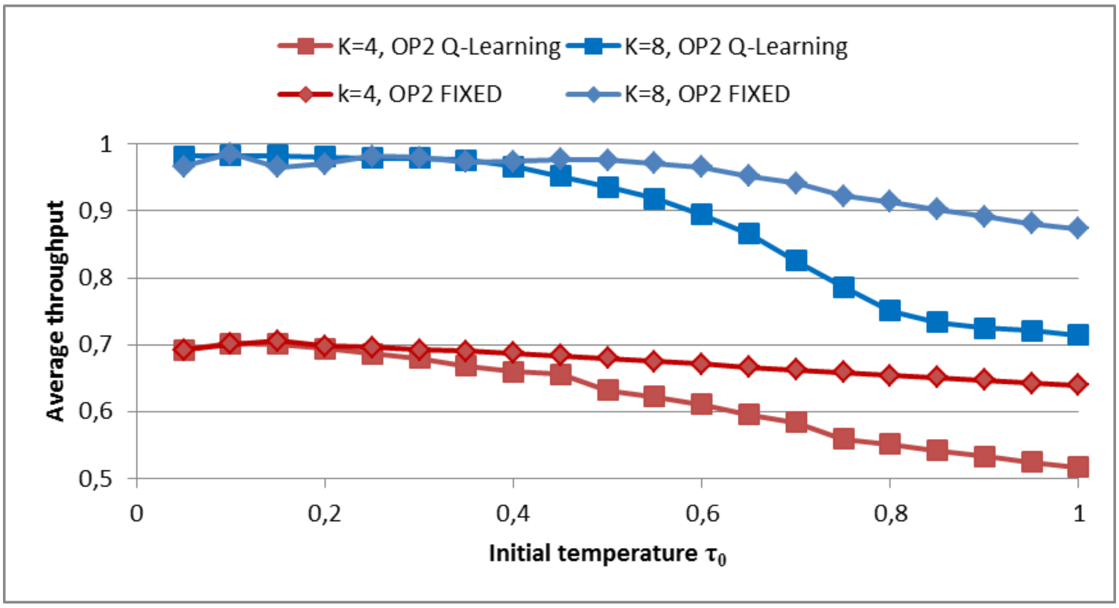


Fig. 6.8 OP1: Q-Learning, OP2: Q-Learning or Fixed for K=4 and K=8

The four curves show the importance that the number of channels has on the average throughput. Regardless of the algorithm operator 2 applies, when there are 8 channels available the performances are much better, at least for τ_0 less than 0.5.

There is a difference of about 0.3 until the temperature gets higher and the curve of Q-Learning with 8 channels declines and approaches the curve of the fixed assignment with 4 channels.

The divergence between Q-Learning and fixed assignment is almost the same regardless the number of channels.

The last analysis in this section assesses the activity time T which indicates the frequency of making channel selection decisions. When T is equal to 1, the operator takes a decision every time step, bringing to better performances, but with the disadvantage of a greater computational load.

In Fig. 6.9 there are the results of a simulation with 8 channels and both operators using Q-Learning. Operator 1 has T=150 in both cases, while Operator 2 has T=150 in one case and T=1 in the second one.

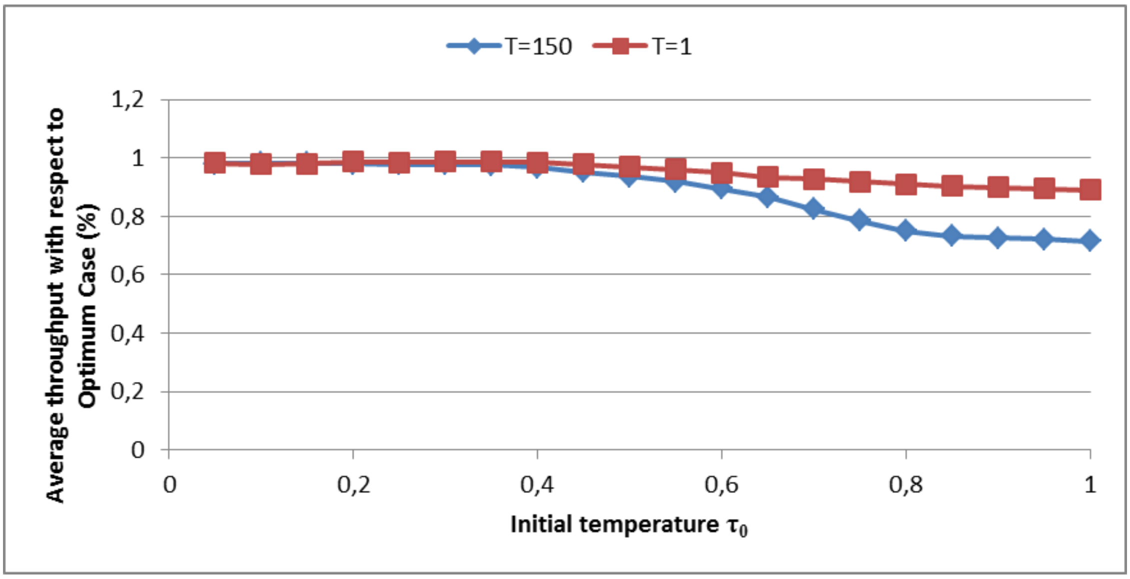


Fig. 6.9 OP1: Q-Learning, OP2: Fixed. Cooling Samples

The two curves have the same behaviour when the initial temperature is lower than 0.5 but the gap between them increases after this value. Throughput is worse if T is greater than 1 because the small cells wait longer before making a new decision, so the last configuration chosen might not be the optimal one for all the duration of the interval between two decisions.

The main results of the analyses conducted in this section can be summarized like this:

- The initial temperature affects directly the average throughput of the system. Low values of τ_0 provide better performances in all of the cases presented, regardless the different configurations of the other parameters.
- Since low temperature causes a greater difference in selection probability for channels that differ in their value function Q(i,k), it is convenient using a cooling function to reduce the value of the initial temperature. When there is no cooling, the Q-Learning strategy does not work well and for high values of τ_0 its behaviour tends to that of a random channel selection. On the contrary, a cooling function improves the overall performances also when the initial temperature is high and, in particular, the *cooling time* function offers better results than the *cooling samples* function.
- When both operators implement the Q-Learning algorithm, the performances are lower than the situation where one of them has a fixed assignment. This is valid in

both the cases of 4 and 8 channels available, even if the results are obviously better in the latter case because every small cell uses its own channel and it does not need to share it.

6.3. Case study 2: Learning Rate

The Learning Rate is one of the most important parameters of the Q-Learning algorithm. It determines the impact of the newly acquired information on the future choices the system will do. This parameter influences the value function Q(i,k) which measures the expected reward that can be achieved by the i-th small cell when it uses the k-th channel. Whenever a channel k is selected by a small cell i, the value of Q(i,k) is updated keeping into account the past experience and the expected reward. The new value of Q(i, k) depends on the Learning Rate according to the formula:

$$Q(i,k) \leftarrow Q(i,k)(1-\alpha_L) + \alpha_L * r(i,k) \qquad (6.3)$$

When α_L is equal to 0, the agent does not learn anything new and the previous information of the value function is kept. On the opposite, if α_L is 1, the algorithm will consider only the most recent information, forgetting what was stored before. This parameter affects the value function Q(i,k) and, consequently, the softmax policy for channel selection which is based on it. According to this policy, the probability that the channel k is selected by the small cell i is:

$$\Pr(i,k) = \frac{e^{\frac{Q(i,k)}{\tau(i)}}}{\sum_{k'=1}^{K} e^{\frac{Q(i,k')}{\tau(i)}}} \qquad (6.4)$$

When there is a high value of the Learning Rate, Q(i,k) is characterized more by the new expected reward r(i,k) and less by the previous stored value. This fact means that the system is more ready to adapt to new changes but it tends to make decisions based on incomplete assumptions. The current situation influences the choice, making the algorithm converge rapidly to a solution which may not be the optimal one because it does not take into account the past history. On the other hand, a low α_L keeps the value function almost constant so the evolution of the probabilities depends primarily on the temperature τ(i). In this case the algorithm converges after a longer time, but it may do to a better solution because more information about the previous configurations of the channels is considered when it takes decisions.

Based on the above considerations, the impact of the Learning Rate on the performances of the algorithm is studied in correlation with the Initial Temperature that was discussed in the previous paragraph. In this way it's possible to point out the reciprocal influence of these two parameters and how they should be properly set in order to enhance the output of the simulation.

This study is divided into 2 sections, each of them analysing the effect of α_L from a different point of view. The first part focuses on the convergence of the algorithm, showing both analytical and qualitative results, meanwhile the second part considers the throughput obtained.

6.3.1. Convergence

The aim of this section is to understand to what extend the learning rate affects the convergence behaviour of the algorithm. The convergence is achieved when the selection probability of one channel is above a determined criterion for all the small cells in the scenario, in this case 0.99. The probabilities are updated every time there is a channel selection and the interval between two of them is defined by the activity period of the small cells, which in this case is equal to 1 so a decision is made every time step.

The simulations for this study run a certain number of realizations of the same experiment. During every realization, the simulation runs until it converges to a proper configuration or until it reaches the maximum number of time steps. Increasing the total number of realizations enhances the accuracy of the simulation, but it increases its duration, so it's necessary to find a good compromise between those two. The following results are obtained executing 1000 realizations of a single experiment which lasts for 10000 time steps.

Eight simulations were run to collect data for this section. Each of them checks the convergence percentage and the convergence time for a fixed value of the learning rate while the initial temperature is varied from 0.05 to 1 with a step of 0.05. The eight curves obtained are plotted in the following two pictures that show clearly how increasing or decreasing the learning rate affects the convergence.

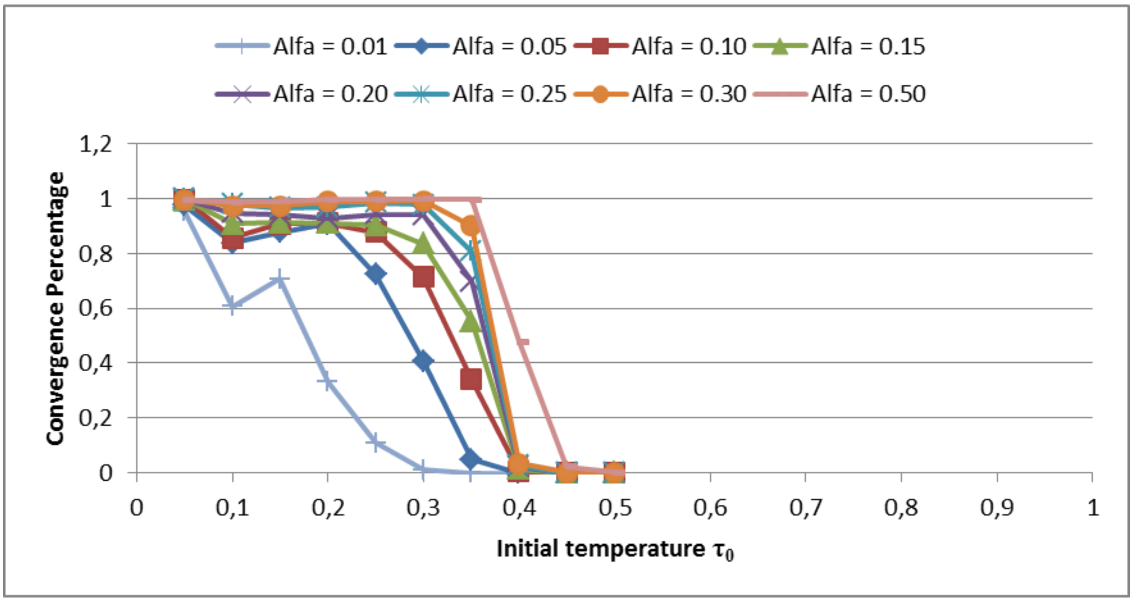


Fig. 6.10 Converge percentage for different values of the learning rate

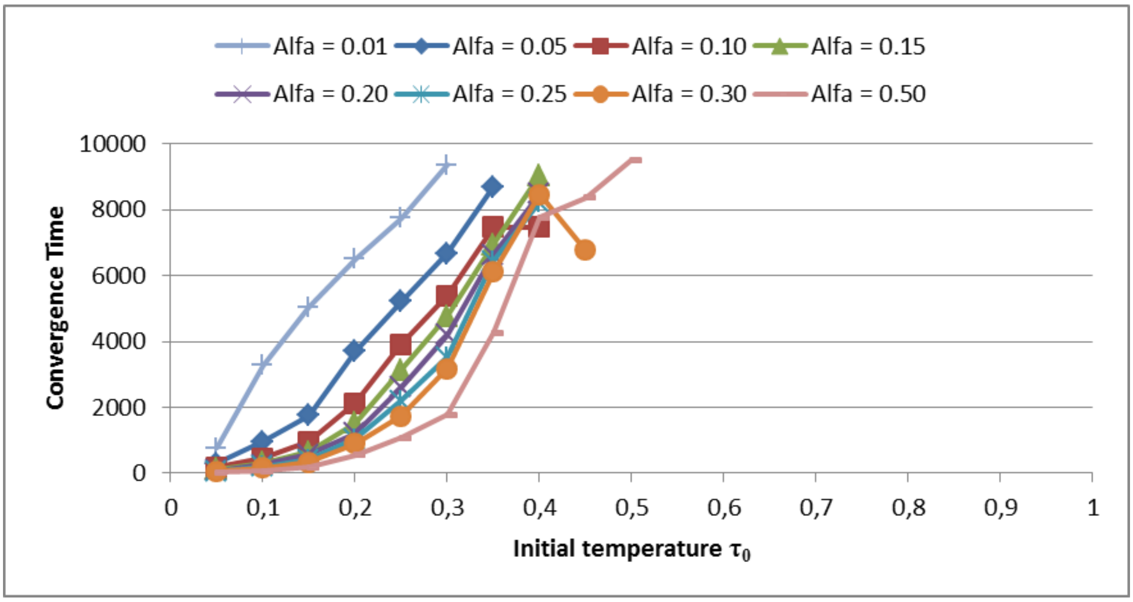


Fig. 6.11 Converge time for different values of the learning rate

The first observation about Fig. 6.10 and 6.11 is that all the curves in both the graphs stop when the initial temperature is equal to 0.5. For values greater than that, the cooling function is not able to reduce the temperature quickly enough to differentiate the selection probabilities before the last time step. Since all the channels are almost equi-probable, the time needed to converge is way longer than the 10000 time steps available, so the algorithm does not converge at all.

The convergence percentage, represented in Fig. 6.10, is equal to 1 for all the values of learning rate when the initial temperature assumes its lowest value. If τ_0 tends to 0, the temperature does the same thanks to the cooling function so the exponential at the numerator of the formula for the selection probability tends to infinite. The probability that every small cell chooses its optimal channel is greatest therefore the algorithm converges in all the cases, regardless the value of α_L. When the initial temperature increases, the value of Q(i,k) and, consequently, that of the learning rate become more relevant for the computation of the selection probability. As a matter of fact, the curves in Fig. 6.10 have better levels of convergence as long as the value of α_L rises.

The convergence time, illustrated in Fig. 6.11, confirms this trend. The algorithm needs less time steps to fully converge when the initial temperature is low. For example, in correspondence of the lowest value of τ_0, the time needed to converge is in the order of one or two hundreds time steps for almost all the values of the learning rate. The worst case is when α_L is equal to 0.01 and the time steps required to converge are 754, which are much more than the other cases but an order of magnitude less compared to the 10000 time steps of the simulation. When the learning rate is so small, the system needs always more time steps to partially converge to a solution, as long as the initial temperature rises. The algorithm converges just for 1.1% after almost 10000 time steps when τ_0 is equal to 0.3 and it does not converge at all for greater values. On the contrary, bigger values of learning rate allow the system to have much higher percentages of convergence and in shorter periods of time.

6.3.2. Qualitative analysis

A qualitative analysis can be conducted by observing the evolution in time of the channel selection probabilities. There is no need to take the average of many experiments because the average throughput is not considered for this part of the study so one experiment is sufficient. Different values of learning rate and initial temperature are tested.

In the challenging case of just 4 channels available for 8 small cells, it is inevitable that two SCs have to use simultaneously the same channel. However, if they do not detect each other, they do not need to share the channel in the time domain so the throughput is higher. The reciprocal distances among the small cells and the propagation conditions of the environment determine the detection range of a SC. When two close small cells share a channel in the time domain, their throughput is reduced to the fixed value 0.5.

With reference to the scenario described in paragraph 6.1, a combination is optimal if all the small cells that share the same channel are distant enough not to detect each other, exception made for SC3 and SC6. Since these two SCs are located in the middle of the scenario and they can detect all the other small cells, they are forced to share a channel in the time domain. A solution is optimal if SC3 and SC6 are paired because in this case all the other SCs can be combined in a way that they do not see each other.

In Tab. 2 is showed one of the possible optimal configurations resulting from the computation. In the first row the small cells are displayed following the order that they have in the reference scenario, while in the second row the corresponding optimal channels are indicated. In this case, SC1 and SC7 share channel 1, SC5 and SC4 share channel 4, SC2 and SC8 share channel 2 and SC3 has to share channel 3 with SC6.

SC	1	5	2	6	3	7	4	8
Channel	1	4	2	3	3	1	4	2

Table 2 Best combination of Small Cells and channels

The first comparison presented in Fig. 6.12 and 6.13 is between α_L equal to 0.01 and 0.15 when the initial temperature is 0.05. In the first case (Fig. 6.12), the best possible configuration is reached after around 600 time steps. Almost all the small cells choose the right channel in a shorter time, but SC3 and SC5 initially do not select the optimal one. After 450 time steps the probability that SC5 uses channel 3 decreases for 150 time steps until it becomes equal to 0.5; SC3 has the same behaviour with reference to channel 4. The optimal configuration is reached when the two small cells switch these equi-probable channels and, since the reward after the change is higher, both the probabilities grow until they reach 1.

In the second case (Fig. 6.13), when the learning rate is higher, the algorithm converges to a solution in a shorter time. After around 20 time steps the channels are already chosen, even if the probabilities of SC2 and SC3 using channel 3 fluctuate for some other time steps. Though, the configuration reached is sub-optimal because even if the small cells are not combined in the best way, some of them do not need to share a channel in the time domain. More precisely, SC1 works with SC4 instead of SC7 and SC5 works with SC8 instead of SC4. These four small cells work fine because they are distant enough from each other but the remaining four, SC2, SC3, SC6 and SC7, are located in the center of the scenario so they all can detect each other.

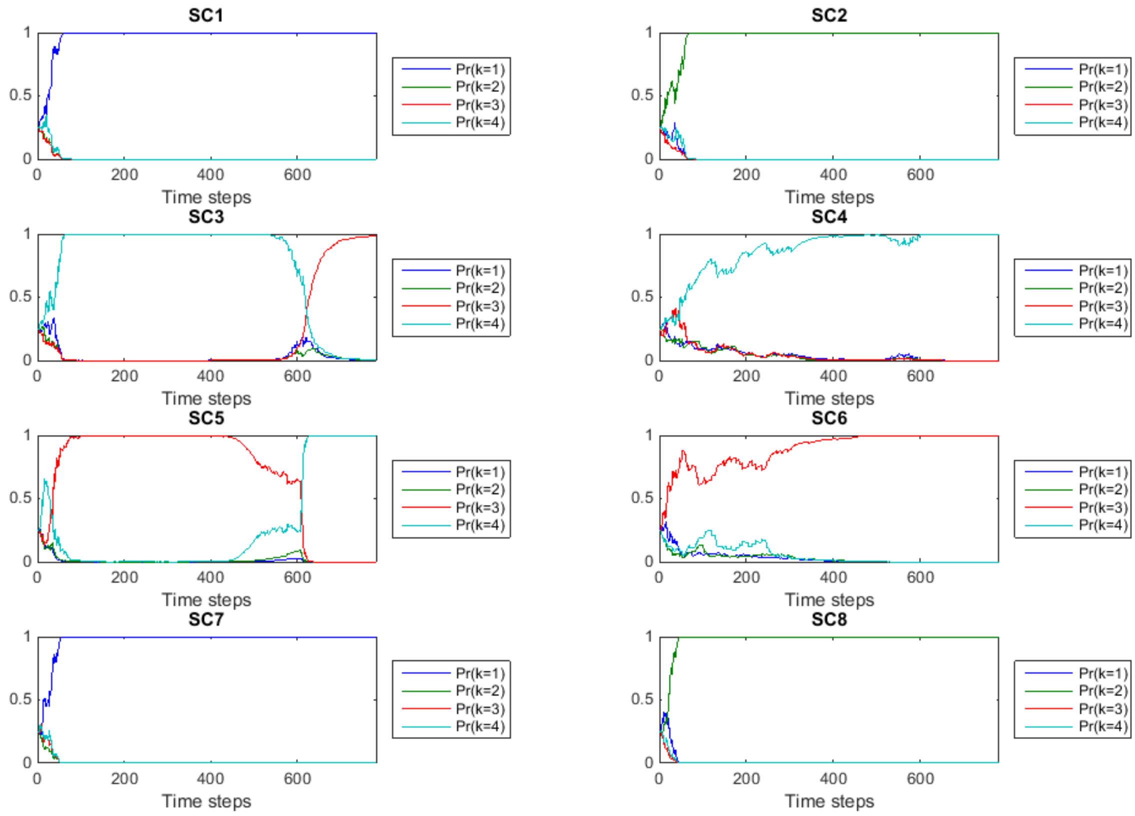


Fig. 6.12 Selection probabilities for τ0 = 0.05 and αL = 0.01

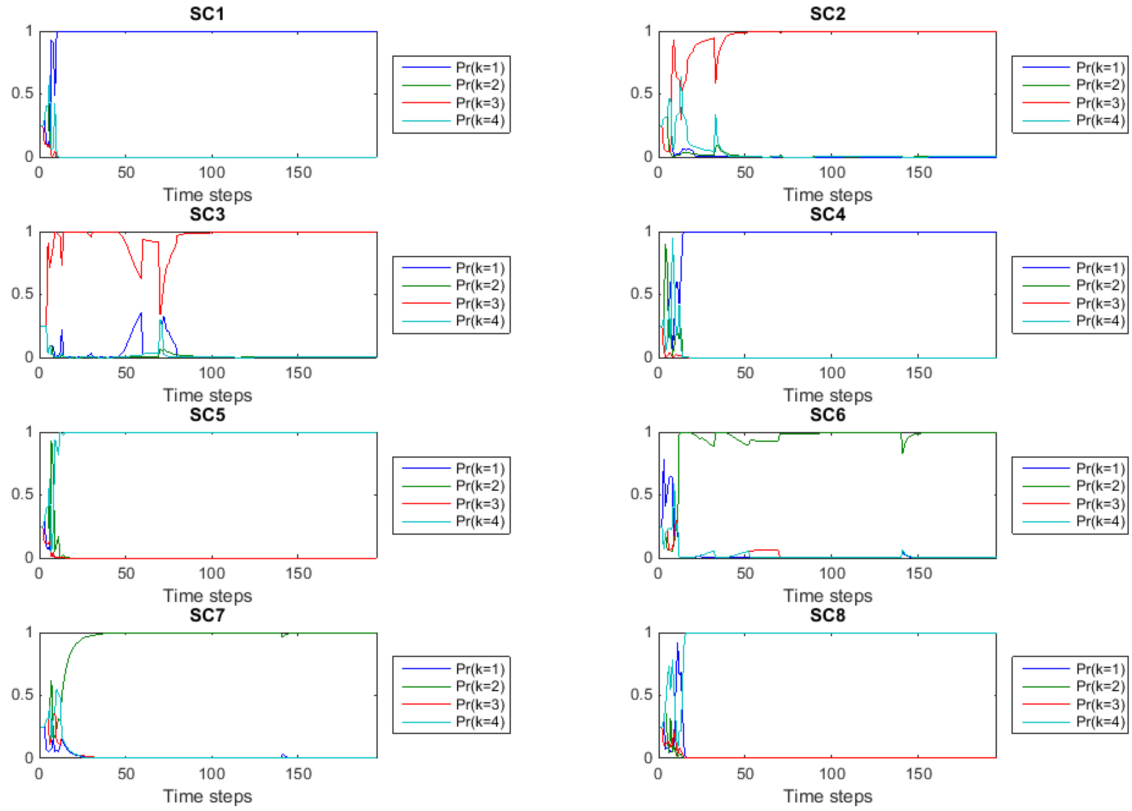


Fig. 6.13 Selection probabilities for τ0 = 0.05 and αL = 0.15

The second comparison presented in Fig. 6.14 and 6.15 is for the same values of the learning rate but for the initial temperature equal to 0.30. When α_L is equal to 0.01, the algorithm converges after less than 500 time steps, except for SC2 and SC3 which both have almost the same probability to select channels 1 and 2, until they take a decision at around the eight hundredth time step. The solution found is not the optimal one because SC2 shares a channel with SC4 instead of SC8 and SC5 shares another channel with SC8 instead of SC4. The other four small cells are instead paired in the best possible configuration.

In the following case, where the learning rate is equal to 0.15, the algorithm converges after a shorter time, like the study about convergence suggested. All the small cells, except for SC3 and SC5, converge after 50 time steps most, which is much less than the previous case. The optimal solution is reached at the one hundredth time step, when SC3 and SC5 switch the channels they have selected at the beginning.

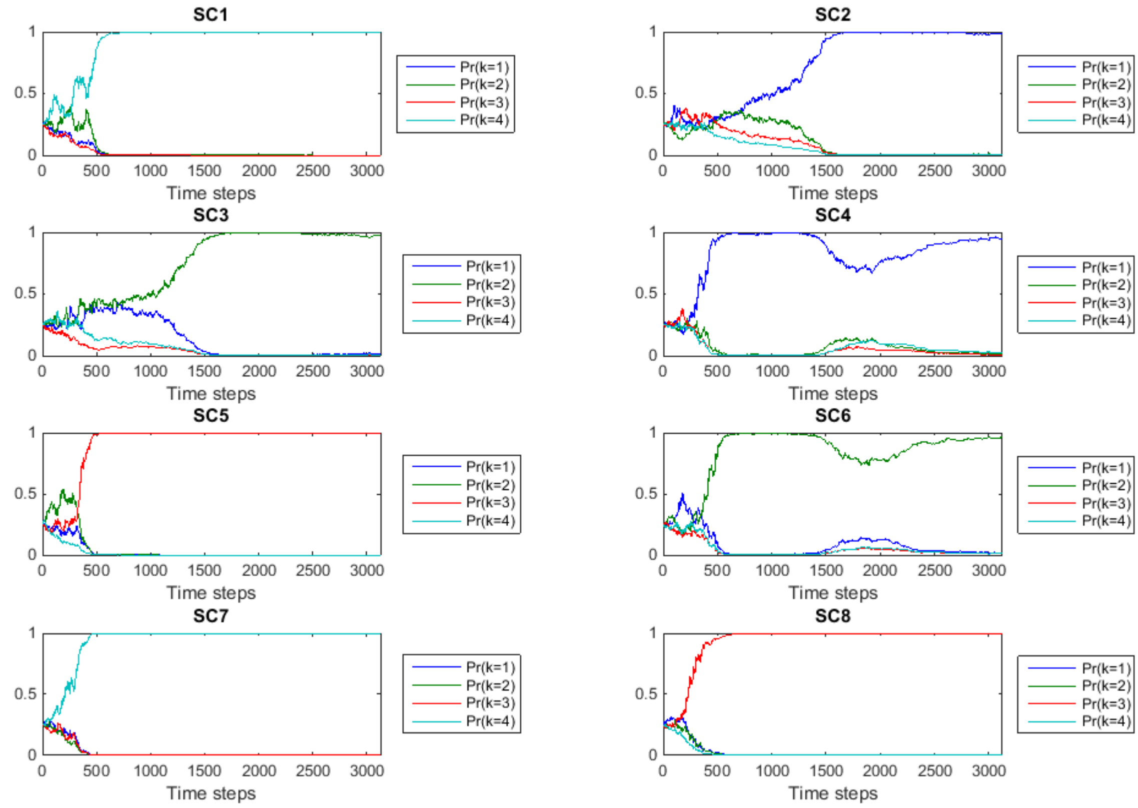


Fig. 6.14 Selection probabilities for τ0 = 0.30 and αL = 0.01

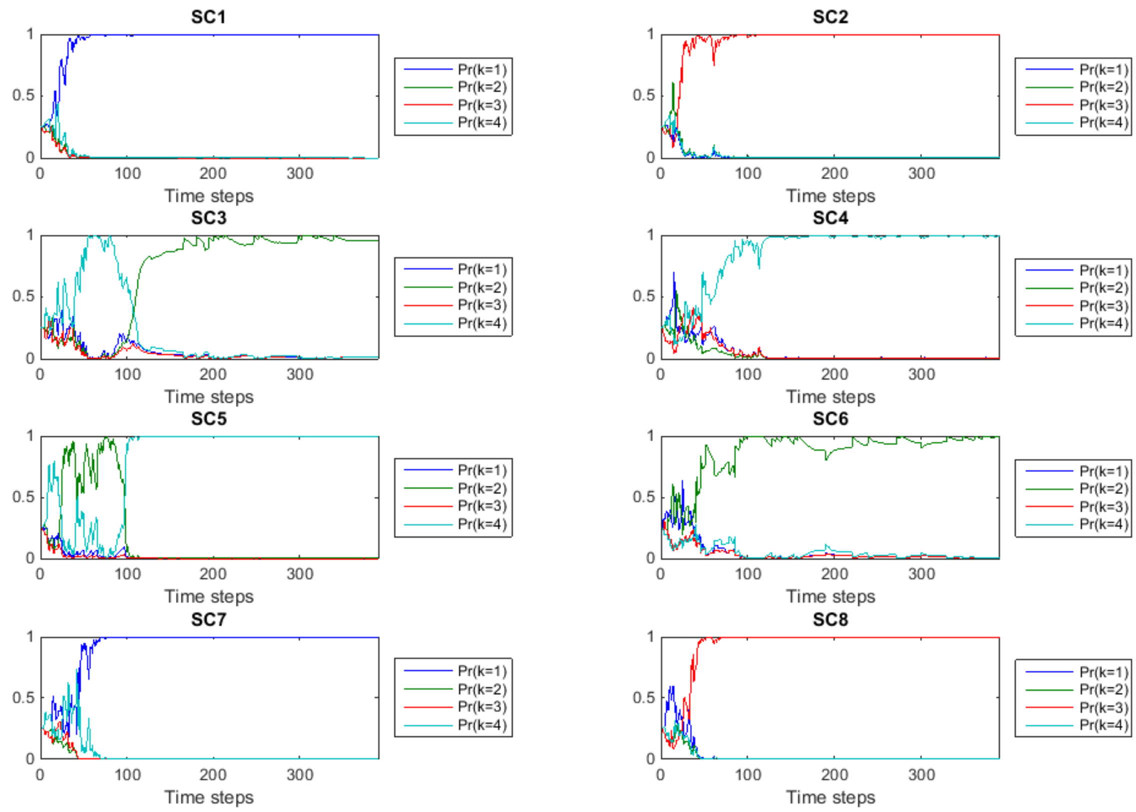


Fig. 6.15 Selection probabilities for τ0 = 0.30 and αL = 0.15

The outcome of this qualitative analysis can be visualized also in Tab. 3, where the normalized throughput is shown for the four cases presented. In this case the throughput is not averaged on the whole duration of the simulation, but it is the aggregated throughput that the 8 small cells obtain in the last time step, after the algorithm has converged to a stable configuration.

Normalized throughput	$\alpha_L = 0.01$	$\alpha_L = 0.15$
$\tau_0 = 0.05$	0,753718	0,66974723
$\tau_0 = 0.30$	0,69643775	0,753718

Table 3 Normalized throughput for different values of τ_0 and α_L

The results obtained in this section show that the effect of the learning rate on the performances cannot be characterized independently from the initial temperature.

6.3.3. Throughput

This section analyses the average throughput obtained with different values of learning rate. Before presenting the results of this study, here is explained the reason why the duration of a simulation has been set to 10000 time steps. This number is large enough to provide an accurate outcome for the range of values of the initial temperature that are significant for this kind of study. Since the research about the convergence performed in the first paragraph presented results only for τ_0 lower than 0.5, it's convenient focus the attention on those values. Obviously, a larger number of time steps seems to bring better results because the throughput obtained is averaged on a longer period and, in this way, the exploratory phase, where the system has not converged yet, has a minor weight on the outcome.

The comparison of the cases with simulation duration of 10000 and 100000 time steps in Fig. 6.16 shows that this effect is noticeable only for high values of the initial temperature. This happens because when τ_0 is low, the exploratory phase is shorter and its reduced throughput does not affect much the average. On the contrary, when the initial temperature is high, the number of time steps needed to converge is comparable to 10000 so the performances are low. If the simulation is longer, the average depends principally from the throughput obtained after the algorithm has reached a stable situation.

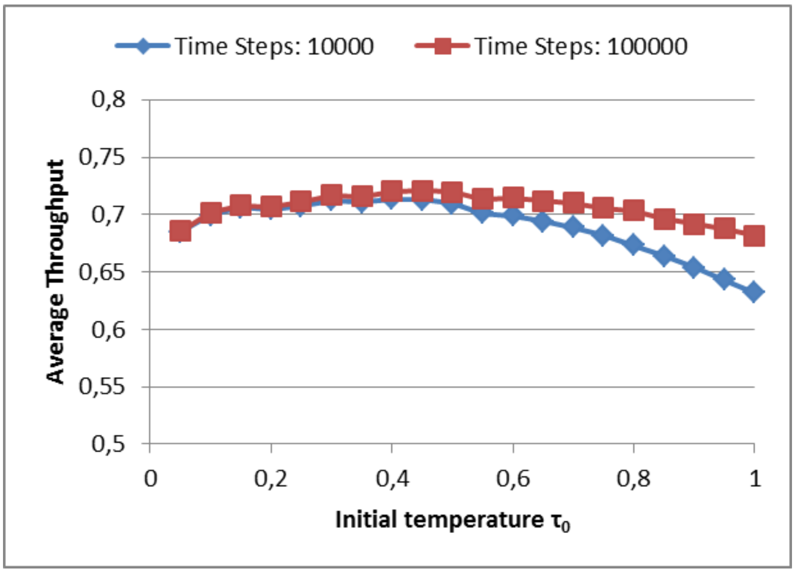


Fig. 6.16 Effect of the number of time steps on the average throughput

The impact of the Learning Rate can be studied by executing some simulations which run 50 experiments of duration 10000 time steps for different values of α_L. The results achieved are plotted in Fig. 6.17 where the average throughput is function of the initial temperature and every curve represents a value of the learning rate.

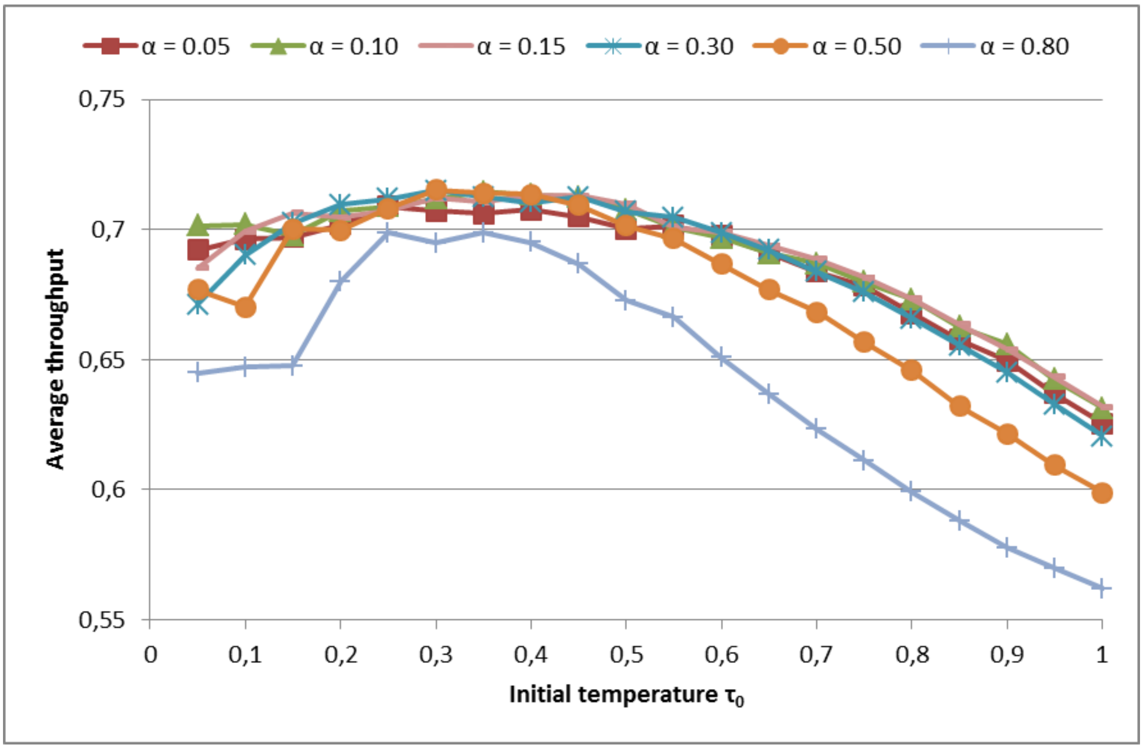


Fig. 6.17 Average throughput for different values of Learning Rate

The performances of the algorithm get worse for any value of learning rate when the initial temperature becomes higher than 0.5, like it was pointed out from the study about the convergence. An interesting behaviour, which is confirmed by the results of the qualitative analysis, is the trend of the curves in the first half of the range of values of τ_0. All of them rise from the starting point until they reach a peak approximately when τ_0 is equal to 0.3 and then they begin their fall till the end with different slopes, depending on the value of α_L.

The effect of the learning rate is noticeable only when this parameter becomes much higher. The curves referred to values of α_L between 0.05 and 0.30 do not present conspicuous differences from each other. They have slightly different values of throughput when τ_0 assumes its lowest values, but then they have an almost identical behaviour during all the simulation. When the learning rate is equal to 0.5, the trend of the curve is similar to the other ones, but the throughput is a bit lower for high initial temperature. The most evident proof of the degradation of the performances in correspondence to the increment of the learning rate is the curve referred to α_L equal to 0.8. The throughput in this case is inferior for every value of τ_0 and, although the trend of this curve is comparable to that of the other curves, there is no intersection with any of them.

The same results can be observed in Fig. 6.18 where the comparison is made for different values of the initial temperature and the average throughput this time is expressed as a function of the learning rate. In addition to the degradation of the performances for higher values of α_L, other aspects can be examined from this graph.

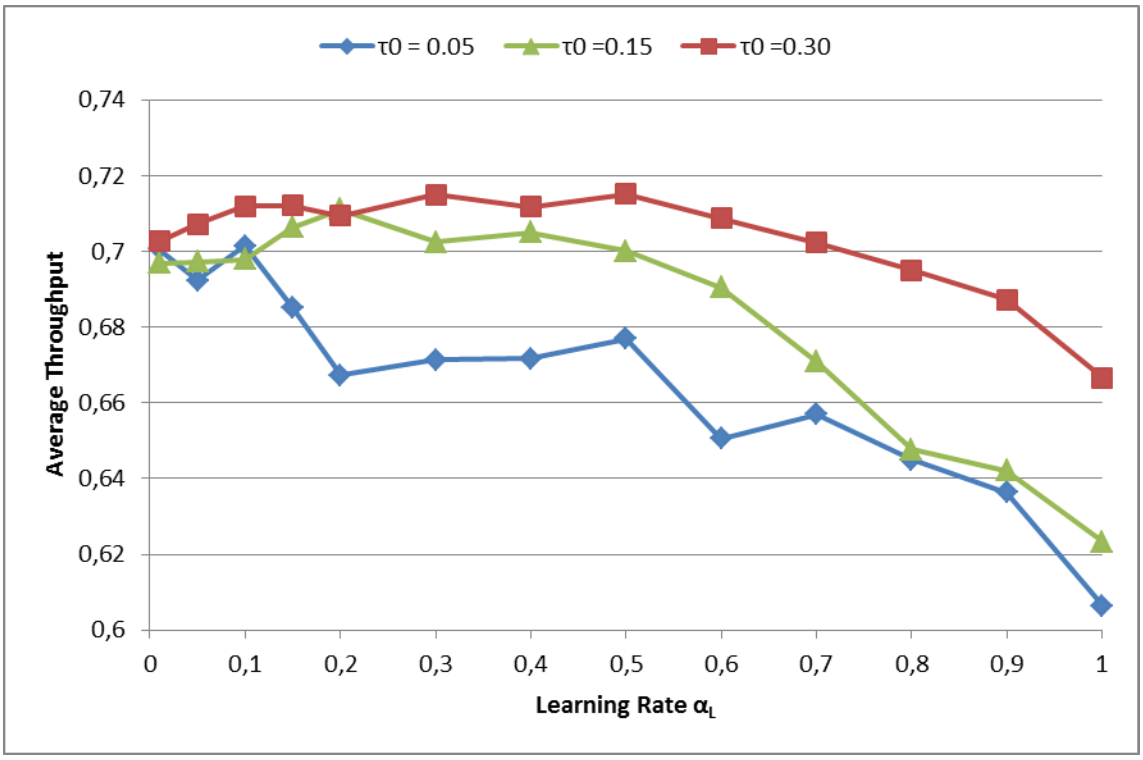


Fig. 6.18 Average throughput for different values of the Initial Temperature

The behaviour of the curves is not the same for all of them, but it depends on the value of the initial temperature. Generally, the throughput is lower for the highest values of the learning rate, but the reduction from the starting value is different case by case. For example, the curve referred to τ_0 equal to 0.05 has not a continuous decrease, but it has a fall, then a modest rise and so on. Even if the trend is similar, the curves referred to τ_0 equal to 0.15 and 0.30 have a slow decline after some slight fluctuations for low values of the learning rate.

In conclusion, the main considerations about the learning rate are:

- The impact of α_L on the behaviour of the Q-Learning strategy is strictly correlated to the initial temperature. The influence of these two parameters on the algorithm is different depending on the type of performance that is considered.

- The convergence analysis showed that the best results are obtained when the initial temperature assumes low values and when the learning rate is high. In this case higher percentages of convergence are reached in a shorter time because the small cells rely more on the newly acquired information from the environment and they take faster decisions about which channel to use.

- The results in terms of average throughput show that different values of learning rate lower than 0.5 offer similar performances, especially when the initial temperature assumes intermediate values around 0.30.

6.4. Case study 3: Modifications in the scenario

The results presented for the previous analyses are based on the simulation scenario described in paragraph 6.1 that was the reference scenario in the 3GPP Study Item [22] and in [1].

This section presents the effects on the performance of the Q-Learning algorithm when some variations are introduced in that scenario. First, the positions and the relative distances and the configuration of the SCs are changed, influencing the propagation losses and, consequently, the detection conditions between them. Then, the number of active users per operator is gradually increased to study how it affects the overall throughput.

6.4.1. Layout 1

In this case, all the SCs are aligned horizontally in the center of the floor building, like in the original configuration, but the distances between the SCs of operator 1 and operator 2 are varied.

The original layout, showed in Fig. 6.1, considers a distance of 5 meters between a small cell of operator 1 and a small cell of operator 2. For this study, the evaluations were made considering a relative distance of 1 meter (Fig. 6.19), 10 meters, and 15 meters (Fig. 6.20). This last configuration presents the maximum possible distance between the SCs, because otherwise SC8 would be out of the floor building. Moreover, the peculiarity of this layout is that every small cell has the same distance, 15 meters, from the previous and the following cell of the other operator.

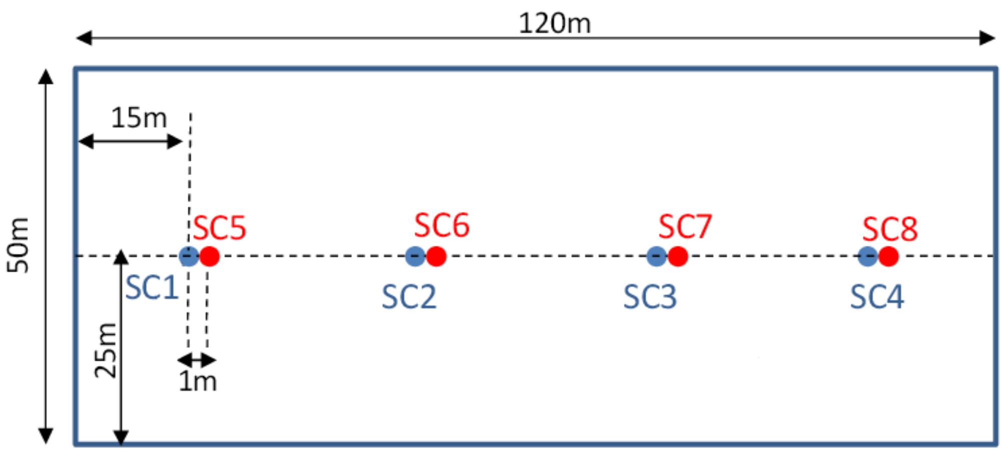


Fig. 6.19 Layout with a distance of 1 m between SCs of different operators

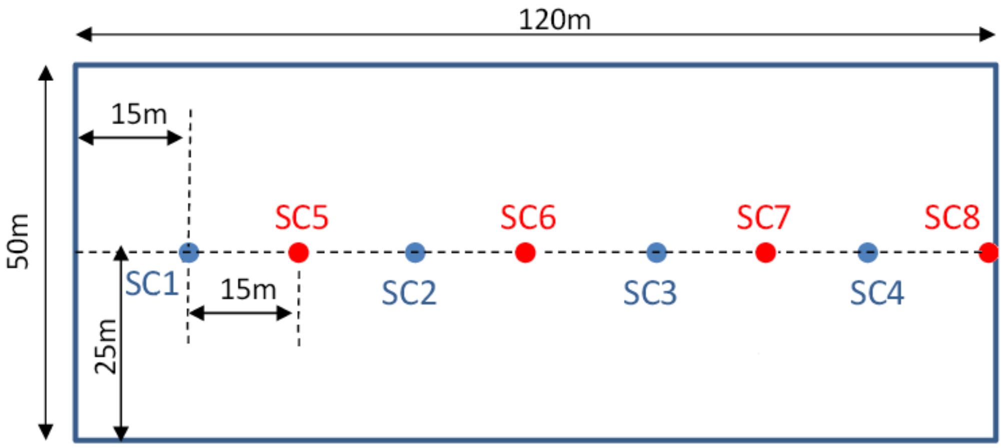


Fig. 6.20 Layout with a distance of 15 m between SCs of different operators

For each of the simulations related to one of these scenarios, the software computed the detection conditions between the APs referring to the same threshold and the same propagation model used in the previous analyses and described in paragraph 6.1. The output of this computation pointed out the detection conditions are the same when the relative distance is 5 meters or more, but they are different when the SCs of the two operators are just 1 meter apart from each other.

More specifically, in the cases where the distance is 5, 10, and 15 meters, only SC3 and SC6 can hear the activities of all the other small cells because they are placed in the middle of the floor building. About the other APS, SC1 is not able to detect the transmissions of SCs 4, 7 and 8; SC2 does not detect SC8; SC4 does not detect SCs 1 and 5; SC5 does not detect SCs 4 and 8; SC7 does not detect SC1; and SC8 does not detect SCs 1, 2 and 5.

On the other hand, in the layout where the distance is of 1 meter, there are 4 central small cells able to detect the transmissions of all the others and they are SC2 and SC3 of operator 1 and SC6 and SC7 of operator 2. On the contrary, the remaining 4 small cells hear in a symmetrical way all the other APs, except for the two farther ones. SC1 and SC5 do not detect the transmissions of SC4 and SC8 and vice versa.

Based on the above considerations, it is expected that the layout with 1 meter distance will offer lower performances because of the higher number of detected APs. In the formula considered for the throughput computation, explained in the paragraph 5.1, the number of APs received above the threshold is a factor at the denominator so, the higher it is, the lower the overall throughput will be.

The simulations executed to study the different scenarios were composed of 200 experiments, each of them with duration 10000 time steps. The parameters of the Q-Learning algorithm are α_L = 0.15, τ_0 = 0.30 and Q_{ini} = 0.5. The results obtained are showed in Fig. 6.21.

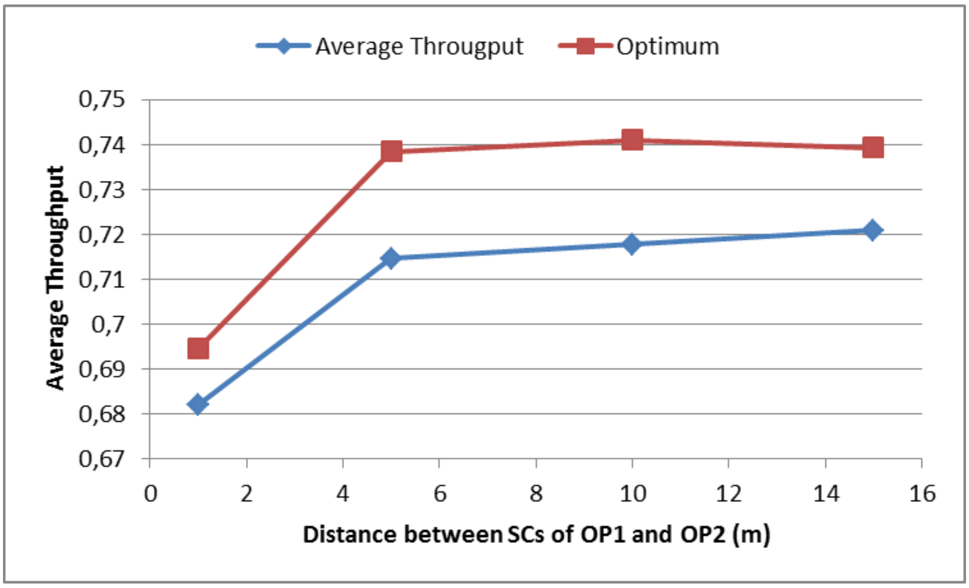


Fig. 6.21 Average throughput as a function of the distance between SCs

The trend of the curves in the picture shows clearly that increasing the distance between the small cells has a beneficial effect on the throughput obtained. The average throughput rises from 0.682 in the case of 1 meter to 0.72 in the case of 15 meters distance. It has the same trend of the optimum throughput achievable, reaching between 96.6% and 98% of it.

It is important to point out that the layout where the distance between small cells is just 1 meter has worse performances compared to the other three cases. The reason of this diversity was discussed earlier and it depends on the different detection conditions. When the distance is 5, 10, and 15 meters, the SCs are able to detect the same number of small cells so the throughput achieved is similar and it is, respectively, 0.714, 0.717, and 0.72. On the contrary, a relative distance of 1 meter presents worse detection conditions which lead to an average throughput of 0.682.

6.4.2. Layout 2

For this part of the study, the positions of all the small cells in the floor building are modified from the original layout. The SCs of the operator 1 are aligned in the bottom half of the scenario, while those of operator 2 are aligned in the top half. SC1 and SC5 are placed at 15 meters from the left side of the layout and the distance between cells of the same operator is still 30 meters, like in the previous cases. This section focuses on what happens when the distance between the two arrays of small cells changes. In particular, four cases are presented: distance equal to 1 meter (Fig. 6.22), 10 meters (Fig. 6.23), 30 meters (Fig. 6.24), and 50 meters (Fig. 6.25) which is the maximum distance achievable because it means placing the small cells at the top edge and bottom edge of the scenario.

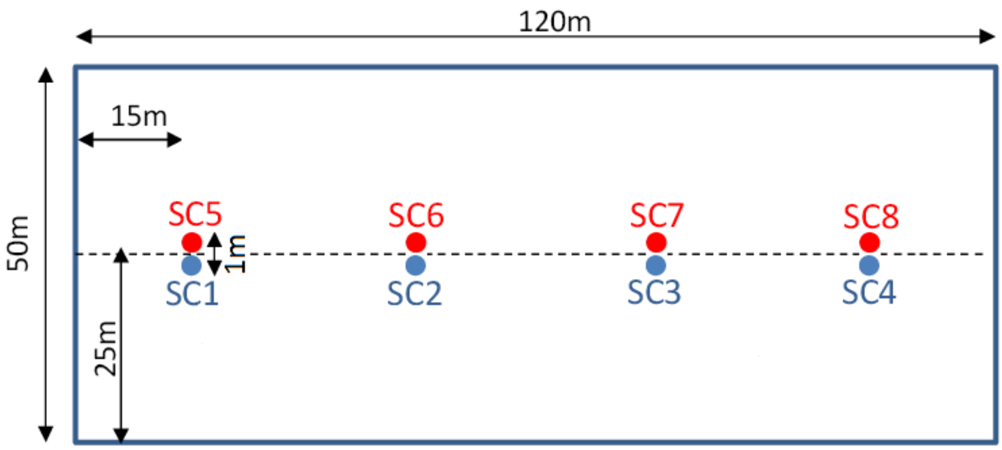


Fig. 6.22 Layout with a distance of 1 m between SCs of different operators

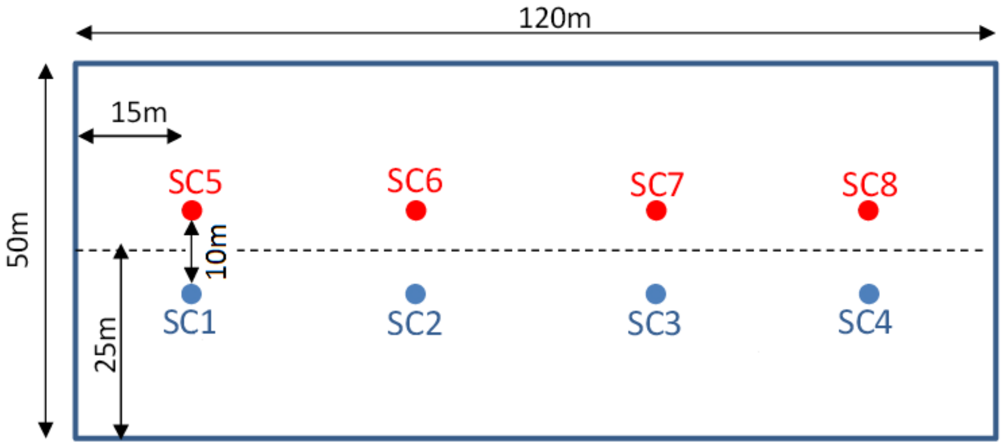


Fig. 6.23 Layout with a distance of 10 m between SCs of different operators

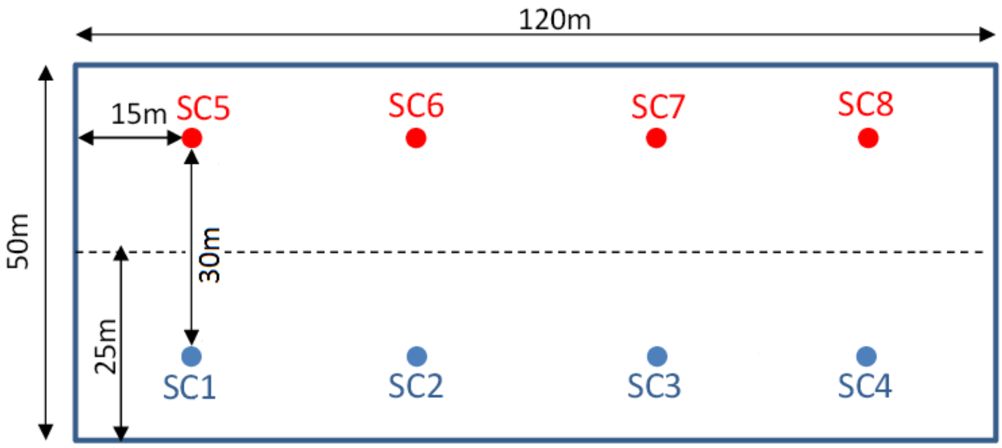


Fig. 6.24 Layout with a distance of 30 m between SCs of different operators

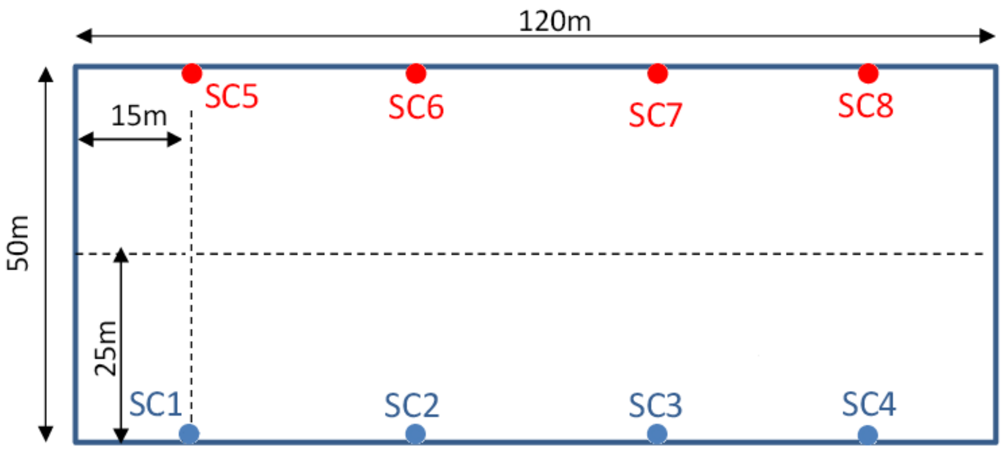


Fig. 6.25 Layout with a distance of 50 m between SCs of different operators

As for the previous study, the software computed the detection conditions for all the four cases presented here, finding that there are two possible configurations, depending on how far the two groups of small cells are.

In the first arrangement, valid for the shortest distances, i.e. 1 meter and 10 meters, SC2, SC3, SC7, and SC8 can detect the transmissions of all the other small cells because they are placed in the middle of the scenario. Each of the four SCs located at the corners instead cannot detect the two small cells on the opposite side of the layout.

On the other hand, when the two arrays are distant enough, i.e. 30 meters and 50 meters, the detection conditions are different and none of the small cells is able to detect the transmission of all the others. The four central SCs cannot hear the small cell placed in the opposite corner, more in detail, SC2 does not detect SC8, SC3 does not detect SC5, SC6 does not detect SC4, and SC7 does not detect SC1. The small cells placed in the corners are not able to hear the last SC of the row and the two SCs in the opposite corner, which belong to the other operator. Specifically, SC1 does not detect SC4, SC7, and SC8; SC4 does not detect SC1, SC5, and SC6; SC5 does not detect SC3, SC4, and SC8; SC8 does not detect SC1, SC2, and SC5.

Also in the cases presented for the previous study on the modifications of the scenario there were two different configurations of the detection conditions. The results of the simulations confirmed that the performances are worse when the SCs are closer because they detect more of the other small cells and this influences negatively the throughput.

Even if these motivations are still valid for the cases presented in this section, there is another factor to be considered here. When the SCs are located near the edges of the floor building, not only the detection conditions change, but also the distance of the users from the serving APs.

At the beginning of every simulation, the users are randomly and uniformly deployed in the scenario, according to a pseudo-random sequence generated by the software from a given seed. The seed chosen for the experiments of this section is the same used in the previous one, so the users are deployed in the same positions as before. While in the previous cases just the small cells of operator 2 were moved by few meters in the middle of the scenario, in these cases all the 8 small cells are moved by many more meters along the whole length of the layout.

The simulations related to this study were run with the same parameters used previously, i.e. 200 experiments, each of them with duration 10000 time steps, α_L = 0.15, τ_0 = 0.30 and $Q_{ini} = 0.5$. The average throughput achieved is depicted in Fig. 6.26.

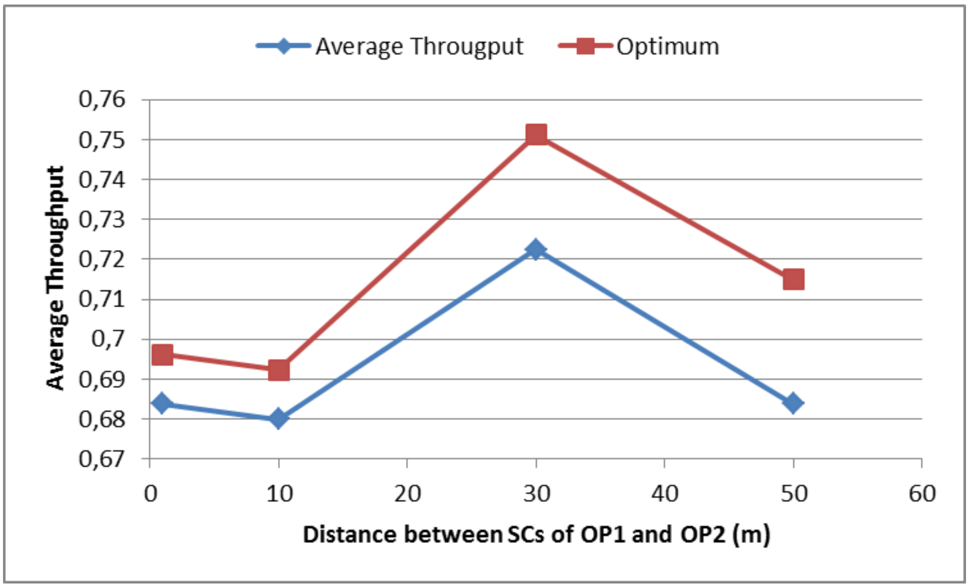


Fig. 6.26 Average throughput as a function of the distance between SCs

The curve in this case does not have a linear trend as a function of the relative distance between the small cells for the reasons aforementioned. While it is expected to have the highest throughput at the distance of 50 meters because the small cells are far from each other, this is not true because of the distance of the users. The SCs do not cover the area in an optimal way because they are placed at the edges of the layout and, moreover, they are grouped according to the operator. This means that if a UE of operator 1 is deployed in the top half of the scenario, it will be very far from its serving AP placed in the bottom edge. This greater distance conduces to a higher interference which decreases the SINR of the n-th user, so the overall throughput is negatively affected.

This effect does not seem to be relevant when the distance between the SCs is 30 meters. In this case the achieved throughput, equal to 0.722, is the highest of the four cases. The users are deployed in such a way that they are close enough to their serving APs so their SINRs are high.

The remaining two cases of 1 meter and 10 meters have worse detection conditions compared to the previous two. This fact is partially supported by the results obtained because the average throughput achieved is worse than the cases with better detection conditions, but the difference with the 50 meters case is slight. The average throughputs of the 1 meter, 10 meters, and 50 meters cases are too similar to deduce a plain interpretation without considering in detail all the factors that may have led to this result.

In any case, the output of the simulation made clear that the effects on the throughput obtained cannot be attributable just on the detection conditions. They may be the main reason for the fluctuation of the performance when different distances between small cells are considered, but they are not the only one. The relative distances of the UEs from the APs have a primary role too in determining the quality of a given layout in terms of the achieved throughput.

6.4.3. Number of active users

This part of the study wants to assess the influence of the number of active users on the average throughput of the system. In all the simulations considered till now, the operators had 10 users each, for a total of 20 UEs in the scenario. For this analysis, the total of users has been modified, increasing and decreasing the number of UEs of both the operators of the same amount. The simulation took place in the default scenario described in paragraph 6.1, and the values of the parameters of the Q-Learning algorithm were $\alpha_L = 0.15$, $\tau_0 = 0.30$ and $Q_{ini} = 0.5$. The number of experiments was 100, with duration 10000 time steps. The results of the simulation are shown in Fig. 6.27.

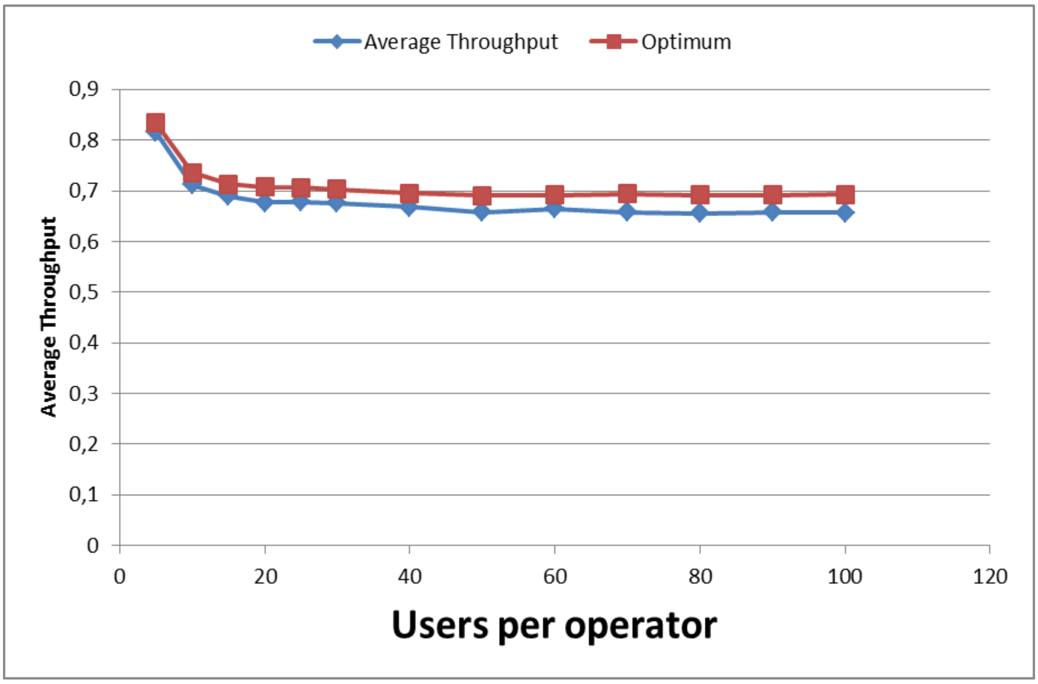


Fig. 6.27 Average throughput as a function of the number of users per operator

The trend of the curve obtained starts from the value of 0.815 when there are just 5 users per operator and then it rapidly declines until it becomes almost constant at 0.70 for any number of active users. The average throughput almost reaches the optimum throughput, especially in the first part of the curve, when the number of UEs per operator is still low.

What emerges from the graph is that the system seems to work better when there are few users in the scenario. Both the total number of users and the SINR that each of them gets are factors in the formula for the computation of the average throughput. If there are few users, generally they have a better SINR since they do not interference much with each other. Therefore, higher SINRs averaged on few users correspond to better performances. On the contrary, when the number of users is higher, they tend to have a lower SINR so the performance normally is worse. However, there the SINR is not determined only by the interference caused by other users, but it depends on the received power, which is conditioned by the position of the UEs from their serving AP. So, when there are many users deployed randomly, some of them will be in bad positions and will interfere with each other, but others will be near the AP and will get a high SINR. The average of the different SINRs made on many users is the reason whereby the performances of the algorithm are very similar when there is a high number of UEs.

In conclusion, the number of active users seems to have little influence on the behaviour of the Q-Learning mechanism which presents performances close to the optimum.

The results obtained from the modifications introduced in the scenario can be summarized like this:

- The distances between the small cells modify the propagation conditions in the scenario which influence the detection conditions. When the distances are smaller, the SCs in the middle of the layout can detect a larger number of SC and this has a negative effect on the overall performances.
- Not only the relative distances between the small cells, but also their positions in the layout have an effect on the average throughput. If the SCs are placed at the edges of the scenario, the distance of the UEs from them is generally increased, producing a reduction of their SINR.
- The results of the simulations suggest that increasing the number of active users beyond a certain threshold does not degrade much the performances of the algorithm.

7. Budget

The research work for this Thesis has been done at the Department of Signal Theory and Communications of the *Universitat Politècnica de Catalunya* (UPC). The equipment utilized is a desktop PC from the laboratory and two servers active 24h per day, provided by University to run long simulations. The piece of software written in Matlab which contained the simulator used for the study was provided to me by the advisor of my Thesis, Professor Jordi Pérez-Romero.

Since the project did not involve tests on field or specific equipment, there were no costs in terms of money, but only in terms of working hours per day.

Without considering the time needed by the servers to run the simulations, some of which took up to 48 hours, the working hours needed for the development of this project were approximately 700. The schedule of the work is presented in the Gantt diagram showed in Table 1 in the Introduction chapter.

After the initial reading about the state of the art of LTE-U and of the Q-Learning algorithm applied to Channel Selection, a certain amount of time was spent to familiarize with the computer program, reproducing some of the results achieved in [1]. Then the studies were conducted sequentially, writing a report of the results obtained at the end of each of them. Meanwhile, the process of searching the contents for the literature review took place, so to write this document in the last two months of work.

8. Conclusions and future development

The work described in this Thesis has concerned with the evaluation of the influence of some key parameters of the Q-Learning strategy on the behaviour of the algorithm within the context of Channel Selection in LTE-U. Understanding to what extent these parameters impact on the overall performance is of considerable importance for exploiting all the potentialities of the algorithm.

A distributed Q-Learning mechanism is the approach proposed in [1] to address the Channel Selection functionality that decides the most appropriate channel in the unlicensed band to set-up a LTE-U carrier for supplemental downlink as a means to facilitate the coexistence. This research is meant to further investigate the characteristics of the Q-Learning algorithm in order to assess the capacity of this approach.

The studies conducted for this project may be summarised as:

- **Initial Temperature**. This analysis showed that the performance of the algorithm improves when the initial temperature assumes low values because in that case there are more differences among the selection probabilities so it is easier for the SCs to select the optimal channel. Different conditions were tested in the other studies presented, but the results confirmed that the system works better when the temperature is low. Besides, the simulations showed that the implementation of a cooling strategy enhances the performances that otherwise would be similar to those obtained with a random channel selection.

- **Learning Rate**. The study conducted in this section showed that this parameter has a high correlation with the Initial Temperature in the computation of the channel selection probabilities. When considering the convergence behaviour, the performances are better for high values of α_L and low values of τ_0 because they provide better converge percentages in a shorter time. On the other hand, when considering the average throughput, lower values of learning rate and higher values of initial temperature offer better results. This output suggests that intermediate values of these parameters can optimize the behaviour of the Q-Learning algorithm, because, even if the small cells need more time to reach the convergence, they converge to better solutions.

- **Modifications in the scenario**. This study pointed out that the relative distances between the small cells influence the detection conditions, resulting in a degradation of the performances when the SCs are too close to each other. Besides, also the positions of the small cells in the scenario affect the behaviour of the algorithm because when they are near the edges of the floor building considered, the distance between them and the UEs are generally high, so the SINR of the users are low and this degrades the average throughput obtained.

- **Number of users**. The simulations conducted for this analysis showed that the number of active users in the scenario has a limited impact on the performances. The average throughput is higher when there are few users and it starts decreasing when the number of UEs increases, but then, beyond a certain threshold, it is steady on a value regardless the number of users.

The results obtained confirmed that the Q-Learning strategy is a valid approach for the Channel Selection mechanism applied to LTE-U. A proper configuration of the parameters of the algorithm, based on the specific conditions of the framework where the Q-Learning model is implemented, can improve the performances and provide better results.

This work can be extended by analysing many other aspects that could not be considered in this Thesis, such as:

- The parameters of the propagation model, or the model itself, could be changed to investigate their impact on the behaviour of the Q-Learning.

- The scenario considered for these analyses was a single floor building, but also an outdoor case could be examined, with all the additional challenges that it brings.

- Possible variations of the Q-Learning technique could be taken into account. Instead of having it completely distributed, some coordination can be considered, for example between the SCs of the same operator.

- The process to update the reward could be based not only on the used channel, but also on what is measured on the other channels, finding a trade-off between performance and complexity.

Bibliography

[1] O. Sallent, J. Pérez-Romero, R. Ferrus and R. Augustì, "Learning-based Coexistence for LTE Operation in Unlicensed Bands," in *IEEE International Conference on Communications (ICC2015) - Workshop on LTE in Unlicensed Bands (LTE-U)*, London, UK, June, 2015.

[2] J. Pérez-Romero, O. Sallent, R. Ferrus and R. Agustí, “A Robustness Analysis of Learning-based Coexistence Mechanisms for LTE-U Operation in Non-Stationary Conditions,” in *IEEE Vehicular Technology Conference (VTC Fall 2015)*, Boston, USA, September, 2015.

[3] J. Pérez-Romero, O. SalIent and R. Agusti, “Enhancing Cellular Coverage through Opportunistic Networks with Learning Mechanisms,” in *GLOBECOM*, 2013.

[4] 3GPP, [Online]. Available: http://www.3gpp.org/about-3gpp.

[5] M. Nohrborg, "3GPP," [Online]. Available: http://www.3gpp.org/technologies/keywords-acronyms/98-lte.

[6] C. Gessner, A. Roessler and M. Kottkamp, UMTS Long Term Evolution (LTE) - Technology Introduction, Rohde-Schwarz, 2012.

[7] E. Dahlman, S. Parkvall, J. Sköld and P. Beming, 3G Evolution, HSPA and LTE for Mobile Broadband.

[8] E. Dahlman, H. Ekstrom, A. Furuskar, J. Karlsson, M. Meyer, S. Parkvall, J. Torsner and M. Wahlqvist, “The Long-Term Evolution of 3G,” Ericsson Review, 2005.

[9] M. Rumney, “3GPP LTE: Introducing Single-Carrier FDMA,” Agilent Technologies.

[10] C. Cox, An Introduction to LTE, 2012.

[11] J. Wannstrom, “3GPP,” June 2013. [Online]. Available: http://www.3gpp.org/technologies/keywords-acronyms/97-lte-advanced.

[12] "LteWorld," [Online]. Available: http://lteworld.org/blog/introduction-relay-nodes-lte-advanced.

[13] M. A. M. Al-Shibly, M. H. Habaebi and J. Chebil, “Carrier Aggreagation in Long Term Evolution-Advanced,” in *Control and System Graduate Research Colloquium (ICSGRC), 2012 IEEE*, July, 2012.

[14] I. Poole, "Radio-Electronics.com," [Online]. Available: http://www.radio-electronics.com/info/cellulartelecomms/lte-long-term-evolution/4g-lte-advanced-carrier-channel-aggregation.php.

[15] R. Hellberg, “Qualcomm,” 5 November 2015. [Online]. Available: https://www.qualcomm.com/news/onq/2015/11/04/how-different-are-lte-u-and-laa.

[16] R. Ratasuk, N. Mangalvedhe and A. Ghosh, “LTE in Unlicensed Spectrum using Licensed-Assisted Access,” in *Globecom 2014 Workshop*, 2014.

[17] D. Flore, “Chairman Summary,” in *3GPP workshop on LTE in unlicensed spectrum*, Sophia Antipolis, France, 2014.

[18] R. Zhang, M. Wang, L. X. Z. Z. Cai, X. Shen and L. Xie, “LTE-Unlicensed: The Future of Spectrum Aggregation for Cellular Networks,” *Wireless Communications, IEEE*, 2015.

[19] A. Brydon, "Unwired Insight," 2015. [Online]. Available:

www.unwiredinsights.com/licensed-assisted-access.

[20] R. Kwan, R. Pazhyannur, J. Seymour, V. Chandrasekhar, S. Saunders, D. Bevan, H. Osman, J. Bradford, J. Robson and K. Konstantinou, “Fair Co-Existence of licensed Assisted Access LTE (LAA-LTE) and Wi-Fi in Unlicensed Spectrum,” in *7th Computer Science and Electronic Engineering Conference (CEEC)*, 2015.

[21] “LTE Forum,” [Online]. Available: http://www.lteuforum.org/.

[22] 3GPP, “Study on Licensed-Assisted Access to Unlicensed Spectrum (Release 13),” in *TR 36.889 V13.0.0*, June, 2015.

[23] R. Sutton and A. Barto, Reinforcement Learning: An Introduction, MIT Press, 1998.

[24] 3GPP, “Further advancements for E-UTRA physical layer aspects,” in *3GPP TR 36.814*, March, 2010.

[25] 3GPP, “Radio Frequency (RF) system scenarios,” in *3GPP TR 36.814*, September, 2014.

[26] ETSI, “ Broadband Radio Access Networks (BRAN): 5 GHz high performance RLAN; Harmonized EN covering the essential requirements of article 3.2 of the R&TTE Directive,” in *EN 301 893*, July, 2014.

Glossary

3GPP: The Third Generation Partnership Project

AP: Access Point

ARQ: Automatic Repeat reQuest

CA: Carrier Aggregation

CCA: Clear Channel Assessment

CC: Component Carriers

CS: Channel Selection

CSAT: Carrier-Sensing Adaptive Transmission

DeNB: Donor eNodeB

DTX: Discontinuous Transmission

DVB: Digital Video Broadcasting

EDGE: Enhanced Data rates for GSM Evolution

eMBMS: enhanced Multimedia Broadcast Multicast Service

eNB: eNodeB

EPC: Evolved Packet Core

E-UTRA: Evolved Universal Terrestrial Radio Access

E-UTRAN: Evolved UMTS Terrestrial Radio Access Network

EPS: Evolved Packet System

FDD: Frequency Division Duplex

FEC: Forward Error Correction

FCC: Federal Communications Commission

FFT: Fast Fourier Transform

GERAN: GSM EDGE Radio Access Network

GSM: Global System for Mobile Communications

IFFT: Inverse Fast Fourier Transform

ISDN: Integrated Services Digital Network

ISI: Inter-symbol Interference

ISM: Industrial, Scientific, and Medical

IMT-Advanced: International Mobile Telecommunications-Advanced

ITEL: Iterative Trial and Error Learning

ITU: International Telecommunications Union

KPI: Key Performance Indicator

LAA: Licensed Assisted Access

LBT: Listen-Before-Talk

LOS: Line of Sight

LTE: Long Term Evolution

LTE-U: LTE in Unlicensed Bands

M2M: Machine-to-machine

MIMO: Multiple-Input Multiple-Output

MBSFN: Multicast-broadcast single-frequency network

NLOS: Non-Line of Sight

OFDM: Orthogonal Frequency-Division Multiplexing

OFDMA: Orthogonal Frequency-Division Multiplexing Access

PCell: Primary Cell

PCS: Physical Carrier Sensing

QoS: Quality of Service

RAN: Radio Access Network

RB: Random Back-Off

RL: Reinforcement Learning

RN: Relay Node

RNC: Radio Network Controller

SAE: System Architecture Evolution

SC: Small Cell

SC-FDMA: Single-Carrier FDMA

SCell: Secondary Cell

SDL: Supplementary DownLink

SINR: Signal To Noise + Interference Ratio

SNR: Signal To Noise Ratio

TD: Temporal Difference

TDD: Time Division Duplex

TPC: Transmit Power Control

TSG-RAN: Technical Specification Group - Radio Access Network

VCS: Virtual Carrier Sensing

VoIP: Voice over Internet Protocol

UMTS: Universal Mobile Telecommunications System

UE: User Equipment

UTRAN: Universal Terrestrial Radio Access Network

W-CDMA: Wideband Code Division Multiple Access